Research towards the effective disruption of reproductive competence in Nile tilapia *Oreochromis niloticus*

A Thesis Submitted for the Degree of

Doctor of Philosophy

By

Yehwa Jin, BSc, MSc

UNIVERSITY *of* STIRLING

Institute of Aquaculture, Faculty of Natural Sciences,

University of Stirling, Stirling, UK

September 2018

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification.

Signature of Candidate: _____

                                           Yehwa Jin

Signature of Principal Supervisor: _____

                                         Prof. Hervé Migaud

Signature of Supervisor: _____

                                         Dr. Andrew Davie

Date: September 2018

Acknowledgement

I would like to express my sincere gratitude to my supervisors Prof. Hervé Migaud and Dr. Andrew Davie for their continuous guidance, invaluable advice, support and encouragements throughout my PhD. I also thank Dr. John B. Taggart for his encouragement and all the precious guidance in the lab particularly with fragment analysis and NGS. In addition, I'm grateful to Dr. David Penman, Prof. James Bron and Dr. Naoki Kabeya who gave me lots of valuable advice, especially for the microinjection system. I also thank Jacquie and Debbie for their guidance and help in the molecular and histology lab and Keith, Brian and Silvere for their continuous help in the Tropical Aquarium. I would also like to thank all the colleagues of Breeding and Physiology research team (recently renamed Breeding and Stock Management Group) and my friends who made me feel like home in Stirling and gave me good memories: I had a good fun in K25 PhD office thanks to Marie, Ben, Lucas, Mikey, Aisha, Kristina, Carol, Aqilah, Yu-Ching, Suleiman, Isah, Michael, Pierre and David. Also, I thank Nui, Dim, Andre, Angela, Taslima, Shankar, Khalid, Simao, Sean, Lynn, Athina, Thomas, Cyril, Winarti, Aom, Thao, Lewis, Sam, Elsbeth, Robyn, Sarah-Louise and John Bostock for being a good companion in my PhD journey. It was my pleasure to work with Swapna. I was lucky to work with Baoshan for the comparative study for mutation screening methods (Chapter 6) in my last part of PhD.

I acknowledge and thank Sun Moon University for the grant and University of Stirling for IMPACT studentship. My warm thanks also goes to Prof. José A. Muñoz-Cueto and Dr. Jose Antonio Paullada Salmerón in Universidad de Cádiz who kindly taught me ISH and helped for ISH analysis (Chapter 5).

My special thanks goes to Dr. Hee Jung Hong, Dr. Park's family, Prof. Sungju Jung, Prof. Myungrae Cho's family and Yoonsol's family who made my days in Stilring joyful and memorable. I would also like to thank my friends, Jeong Myeong, Gahee, Yeseung, Jeong un, Sangmin, Prof. Joon Yeong Kwon and Bio-fam members who have endlessly encouraged me. I'm deeply grateful to my beloved Adam for his encouragement and believing in me. The last but not least, my greatest thanks go to my parents and family for their constant support and love, throughout this journey.

Published Articles and Conferences

<u>PUBLISHED PAPERS IN PEER REVIEWED JOURNALS</u>

- **Jin YH.,** Davie A., Migaud H. 2019. Expression pattern of *nanos, piwil, dnd, vasa* and *pum* genes during ontogenic development in Nile tilapia *Oreochromis niloticus.* Gene 688, 62–70.

<u>CONFERENCE POSTER PRESENTATIONS</u>

- **Jin YH.,** Davie A., Migaud H. 2018. Impact of *piwil2* knockout in primordial germ cells using CRISPR/Cas9 system in Nile tilapia *Oreochromis niloticus*. 11th International Symposium on Reproductive Physiology of Fish (ISRPF), Manaus, Brazil, 03-08 June.
- **Jin YH**., Davie A., Migaud H. 2016. Impact of high temperature on reproductive competence of male Nile tilapia *Oreochromis niloticus*. 40th Aquaculture Europe, Edinburgh, Scotland, 20-23 September.
- **Jin YH**., Davie A., Migaud H. 2016. Expression pattern of *nanos* and *piwil* genes during ontogenic development in Nile tilapia *Oreochromis niloticus*. 28th Conference of European Comparative Endocrinologists (CECE), Leuven, Belgium, 21-25 August.

<u>CONFERENCE ORAL PRESENTATIONS</u>

- **Jin YH.,** Davie A., Migaud H. 2018. Impact of *piwil2* knockout in primordial germ cells using CRISPR/Cas9 system in Nile tilapia. 5th Institute of Aquaculture PhD Research Conference, University of Stirling, Stirling, UK, 17 April.
- **Jin YH**., Davie A., Migaud H. 2016. Expression pattern of *nanos* and *piwil* genes during ontogenic development in Nile tilapia *Oreochromis niloticus*. Young investigator symposium, 28th CECE, Leuven, Belgium, 21-25 August.

Abstract

Reproductive containment in farmed fish is highly desired for sustainable aquaculture to prevent genetic introgression with wild conspecifics and enhance productivity by suppressing sexual maturation. A number of strategies have already been implemented or have been tested in commercially important fish (*e.g.* triploidy, monosexing, hormonal therapies); however, they either do not result in 100% containment, or they cannot be applied to all species. One promising new approach consists in disrupting primordial germ cells (PGCs), at the origin of germline cells, to induce sterility. The work carried out in this doctoral thesis aimed to investigate the genes involved in the survival of germ cells and subsequently conduct a functional analysis of candidate genes using CRISPR/Cas9 gene editing system to ultimately provide the basis for the development of a novel sterilisation technique. Nile tilapia was chosen as the experimental animal as it is a major aquaculture species worldwide and the control of reproduction plays a critical role in the farming productivity in this species. In addition, the species has clear advantages as its whole genome sequence is accessible, the generation time is relatively short and zygotes can be available all year round. Initially, a panel of 11 candidate genes with reported roles in survival of PGCs was investigated during the ontogenic development which led to the selection of *piwi-like (piwil)* gene as a target for genome editing. Then, high temperature was tested as a means to induce germ cell loss to better understand the mechanism underlying germ cell survival and apoptosis, and this study confirmed the functional importance of *piwil* genes in relation to germ cell loss and proliferation. In addition, the study suggested potential subfunctionalisation within the Bcl-2 gene family which requires further investigation. The next step aimed to optimise the CRISPR/Cas9 gene editing method by improving the microinjection system and testing different concentrations of sgRNAs. Over 95% of injected embryos showed on-target mutation in *piwil2* via zygote injection of CRISPR/Cas9 reagents and complete KO larvae were shown in half of the mutants, producing putative sterile fish. However, there was no clear association between the phenotypes in PGCs and the mutation rate. Further comparative studies of mutant screening methods including T7E1, RGEN, HRMA, fragment analysis and NGS revealed that the genotypes of F0 are highly mosaic, suggesting that deep sequencing is recommended for accurate and high throughput F0 screening and further improvement for predictable genome editing is required for a reliable gene functional analysis in F0. In summary, the current thesis provided new

scientific knowledge and supporting evidence for the use of the CRISPR/Cas9 gene editing platform to study gene function associated with sterility, with the ultimate goal to develop an alternative sterilisation method in fish.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

11-KT — 11-ketotestosterone

AAVs — adeno-associated viral vectors

A_{diff} — type A differentiated spermatogonia

ANOVA — analysis of variance

A_{und} — type A undifferentiated spermatogonia

BV — blood vessel

Cd205 — cluster of differentiation 205, lymphocyte antigen 75

CPPs — non-viral carrier such as peptide-based carrier

CRISPR/Cas9 — clustered regularly interspaced short palindromic repeats/CRISPR associated effector 9

crRNA — CRISPR RNA

dah — days after hatching

DD — degree-day

DIG — digoxigenin

dpc — days post coitum

dpf — days post fertilisation

DSB — double-strand breaks

E2 — estradiol-17β

ED — efferent duct

EtBr — ethidium bromide

eGFP — enhanced green fluorescence protein

EIA — enzyme immune assay

ELISA — enzyme-linked immunosorbent assay

FADD — Fas-associated death domain

Fas — death receptor

FDA — the US Food and Drug Administration

Fragment analysis — fluorescent PCR capillary gel electrophoresis

FSH — follicle stimulating hormone

G0 — Gap 0 phase, the quiescent (G0) phase of the cell cycle

G1 — Gap 1 phase of the cell cycle, the cells synthesize mRNA and proteins prior to DNA synthesis in S phase

G2 — Gap 2 phase of the cell cycle, the cells grow and synthesize protein prior to mitosis

GABA — γ-aminobutyric acid

Gdf9 — growth differentiation factor

gDNA — genomic DNA

GFP — green fluorescence protein

GM — genetically modified; genetic modification

GMO — genetically modified organism

GMSC — Genetic Modification Safety Committee (University of Stirling)

GMT — genetically male tilapia

GnIH — gonadotropin-inhibitory hormone

GnRH — gonadotropin releasing hormone

GP — germ plasm

GPCR — seven transmembrane G protein coupled receptor

GSI — gonadal somatic index

HCG — human chorionic gonadotropin

HDR — homology-directed repair

HEPN — higher eukaryotes and prokaryotes nucleotide-binding

hpf — hours post fertilisation

HPLC — high performance liquid chromatography

HRMA — high resolution melt analysis

HSI — hepatosomatic index

Indel — insertion or deletion

IP — intellectual property

ISH — *in situ* hybridisation

KO — knockout

LH — luteinizing hormone

MBT — midblastula transition

miRNA — microRNA

MSC — maternal sterility construct

MST — mixed sex tilapia

MT — 17α-methyltestosterone

MZT — maternal to zygotic transition

NGS — next generation sequencing

NHEJ — non-homologous end joining

NILV — non-integrating lentiviral vector

ORF — open reading frame

PAGE — polyacrylamide gel electrophoresis

PAM — protospacer adjacent motif

PGC — primordial germ cell

piRNA — piwi-interacting RNA

Piwi — P-element induced wimpy testis

pPGCs — presumptive PGCs

Puf — Pumilio and FBF, fem-3 mRNA-binding factor

qRT-PCR — quantitative reverse transcription PCR

RED — restriction enzyme digestion

RGEN — CRISPR/Cas-derived RNA-guided engineered nucleases

RISC — RNA-induced silencing complexes

RNP — ribonucleoprotein

RT-PCR — reverse transcription PCR

S — synthesis phase of the cell cycle

sgRNA — single guide RNA

shRNA — short hairpin RNA

sncRNA — small non-coding RNA

SPC — steroid-producing cell

SRT — sex reversed tilapia

SSC — spermatogonial stem cell

SSTNE — Spermidine-Spermine-Tris-NaCl-EGTA

T — testosterone

T7E1 — T7 Endonuclease 1

TALE — transcription activator-like effector

TALEN — transcription activator-like effector nuclease

tracrRNA — trans-activating CRISPR RNA

UTR — untranslated region

VSI — visceral somatic index

WT — wild type

ZFD — zinc-finger domain

ZFN — zinc-finger nuclease

List of species

African Clawed Frog (*Xenopus laevis*)

Amazon catfish (*Pseudoplatystoma corruscans*)

Amazon molly (*Poecilia formosa*)

Ameca (*Ameca splendens*)

Argentinian silverside (*Odontesthes bonariensis*)

Atlantic cod (*Gadus morhua*)

Atlantic salmon (*Salmo salar*)

Barfin flounder (*Verasper moseri*)

Blue drum (*Nibea mitsukurii*)

Blue tilapia (*Oreochromis aureus*)

Brown-marbled grouper (*Epinephelus fuscoguttatus*)

Cachara (*Pseudoplatystoma reticulatum*)

Catarina scallop (*Argopecten ventricosus*)

Channel catfish (*Ictalurus punctatus*)

Chinese sturgeon (*Acipenser sinensis*)

Chinook salmon (*Oncorhynchus tshawytscha*)

Common carp (*Cyprinus carpio*)

Cynomolgus monkeys (*Macaca fascicularis*)

Dusky grouper (*Epinephelus marginatus*)

Eastern mosquitofish (*Gambusia holbrooki*)

European plaice (*Pleuronectes platessa*)

European Seabass (*Dicentrarchus labrax*)

Fruit fly (*Drosophila melanogaster*)

Fugu (*Takifugu rubripe*)

Gibel carp (*Carassius auratus gibelio*)

Goldfish (*Carassius auratus*)

Grass carp (*Ctenopharyngodon idella*)

Grass puffer (*Takifugu niphobles*)

Grey mullet (*Mugil cephalus*)

Guppy (*Poecilia reticulata*)

Half-smooth tongue sole (*Cynoglossus semilaevis*)

Halibut (*Hippoglossus hippoglossus*)

Ice goby (*Leucopsarion petersii*)

Killifish (*Fundulus heteroclitus*)

Largemouth bass (*Micropterus salmoides*)

Loach (*Misgurnus anguillicaudatus*)

Medaka (*Oryzias latipes*)

Mouse (*Mus musculus*)

Mozambique tilapia (*Oreochromis mossambicus*)

Mummichog (*Fundulus heteroclitus*)

Nibe croaker (*Nibea mitsukurii*)

Nile Tilapia (*Oreochromis niloticus*)

Northern pike (*Esox lucius*)

Ocean pout (*Zoarces americanus*)

Orange-spotted grouper (*Epinephelus coioides*)

Pacu (*Piaractus mesopotamicus*)

Patagonian freshwater pejerrey (*Patagonina hatcheri*)

Patagonian pejerrey (*Odontesthes hatcheri*)

Pejerrey (*Odontesthes bonariensis*)

Pintado (*Pseudoplatystoma corruscans*)

Pufferfish (*Takifugu rubripes*)

Rainbow trout (*Oncorhynchus mykiss*)

Red sea bream (*Pagrus major*)

Rohu (*Labeo rohita*)

Rosy barb (*Barbus conchonius*)

Roundworm (*Caenorhabditis elegans*)

Senegal sole (*Solea senegalensis*)

Silver barb (*Puntius gonionotus*)

Southern catfish (*Silurus meridionalis*)

Turbot (*Scophthalmus maximus*)

White croaker (*Pennahia argentata*)

Yellowtail kingfish (*Seriola lalandi*)

Yellowtail tetra (*Astyanax altiparanae*)

Zebrafish (*Danio rerio*)

List of genes

11β-HSD (*11β-Hydroxysteroid dehydrogenase*)

aldh1a2 (*aldehyde dehydrogenase 1 family member a2*)

amh (*anti-Müllerian hormone*)

amhrII (*anti-Müllerian hormone receptor type II*)

amhy (*Y chromosome-specific amh*)

Apaf-1 (apoptotic peptidase activating factor-1)

bax (*B-cell lymphoma 2 associated X*)

bcl-2 (*B-cell lymphoma 2*)

bcl-xL (*B-cell lymphoma-extra large*)

bim (*Bcl-2 homology 3-only protein*)

BMPs (bone morphogenetic protein)

boule (*Deleted in Azoospermia family of RNA binding protein*)

box (*Bcl-2-related ovarian killer*)

cxcl12a (*C-X-C motif ligand 12a*, previously named *stromal cell derived factor 1a*, *Sdf1a*)

cxcr4b (*C-X-C motif receptor 4b*)

cyp19a1 (*cytochrome P450 family 19 subfamily A member 1*)

cyp26a1 (*cytochrome P450 family 26 subfamily A member 1*)

daz (*Deleted in Azoospermia*)

dazl (*daz-like*)

dnd (*dead end*)

dmrt (*doublesex- and mab3-related transcription factor*)

dmy (*DM-domain gene*)

figla (*factor in the germ line, alpha*)

foxl2 (*forkhead box l2*)

gfra1 (*glial cell line-derived neurotrophic factor (GDNF) family receptor alpha-1*)

gsdf (*gonadal soma derived factor*)

hsp70 (*heat-shock protein 70*)

lhcgr (*luteinizing hormone/choriogonadotropin receptor*)

mcl-1 (*myeloid cell leukemia 1*)

Mili (murine homolog of *piwil2*)

Miwi (murine homolog of *piwil1*)

Miwi2 (murine homolog of *piwil4*)

MVH (mouse Vasa homolog)

nanos (RNA binding zinc finger proteins)

ngn3 (*neurogenin3*)

piwil (*piwil-like*)

pum (*pumilio*)

sf-1 (*steroidogenic factor-1*)

sox9a (*sex determining region Y-box 9a*)

sry (*sex determining region Y*)

tdrd7 (*tudor domain containing protein 7*)

vasa (ATP-dependent RNA helicase of the DEAD-box family)

wnt4 (*wingless-type MMTV integration site family, member 4*)

*zili (*zebrafish homolog of *piwil2)*

ziwi (zebrafish homolog of *piwil1*)

CHAPTER 1

GENERAL INTRODUCTION

1.1 Sterilisation methods in fish

Aquaculture plays a crucial role to meet the growing demand for food as well as easing the pressure on wild fish stocks (HLPE, 2014). Whilst the world production from capture fisheries has plateaued since the 1990s, the aquaculture sector has shown a continual growth, contributing to the increase of the world's fish production from 6 kg/capita/year in 1950 to 19.2 kg/capita/year in 2012 (HLPE, 2014). As the world population is growing, reaching 7.4 billion in 2015 and is expected to reach over 10 billion by 2055 (United Nations, 2017), the dependence on the fish supply from aquaculture will continue to increase, which requires extensive research to improve production.

An essential element of the sustainable intensification in aquaculture is the reliable control of fish reproduction. The maturation process of fish often causes negative effects on the productivity and efficiency in fish farming by reducing growth rate and flesh quality (Zohar, 1989). In addition, there is a potential threat to wild fish populations in the case of unintentional escape of farmed fish as they will compete for resources and furthermore there is a perceived risk of genetic introgression through the breeding with wild conspecifics (Muir and Howard, 1999). In some situations, there is a concern related to the spreading of reproductively competent non-native invasive species, and as the boundaries of biotechnology are advanced, the perceived ecological risks associated with genetically modified (GM) organisms must be understood and mitigated (Hu *et al*., 2007). To solve these problems, effective sterilisation of farmed fish is highly desired. Sterile farmed fish that escape from farming systems cannot breed; therefore, hybridisation with wild fish populations and potential introgression can be prevented (Wong and Zohar, 2015a). In addition, by using sterile farmed fish, fish growth rate can be increased by suppressing gonadal development (Ali and Rao, 1989).

Tilapia is the second most farmed fish after carp spp., with the global aquaculture production of Nile tilapia (*Oreochromis niloticu*s) reaching 3.9 million tonnes in 2015 with a sale value of USD 6 billion (FAO, 2017). The farming of all male stocks is preferred as males grow faster than female tilapia and in the case of mixed sex production, maturation is reached prior to attaining harvest weight and the subsequent repeated spawning results in significantly reduced harvest biomass associated with a slow growth rate due to overcrowding and increase of small sized fish (Beardmore *et al*., 2001; Phelps and Popma, 2000). However, tilapia farming largely relies on the rearing of monosex male populations generated primarily through direct hormonal treatment (17α-

methyltestosterone, MT) (Rakocy, 2005), which causes safety concerns. The utilisation of YY supermales to produce genetically male tilapia (GMT) is commercially available as well as showing higher performance than sex reversed tilapia (SRT) or mixed sex tilapia (MST) (Beardmore *et al*., 2001), but it requires intensive broodstock management which hinders wide usage. More importantly, neither methods can prevent the negative impact on the environment caused by escapees. Therefore, a breakthrough in tilapia farming could be achieved by developing protocols to produce sexually sterile tilapia.

To date, various methods to mitigate or eliminate reproduction in fish have been developed or are in development including chromosome set manipulation, hybridisation, gamma irradiation, steroid treatment, chemical treatment, surgical gonadectomy and transgenic sterilisation, but most of them are expensive, not always fully reliable, and can be associated with food safety concerns (Donaldson *et al*., 1993; Kapuscinski and Patronski, 2005). Therefore, a safe, efficient and affordable means to sterilise fish is highly desired.

1.1.1 Non-genetically modified (GM) sterilisation methods

Chromosome set manipulation is a commonly used sterilisation method for some aquaculture species with relatively low cost including Atlantic salmon (*Salmo salar*), rainbow trout (*Oncorhynchus mykiss*) and grass carp (*Ctenopharyngodon idella*) (Arai, 2001; Johnstone *et al*., 1991; Kapuscinski and Patronski, 2005; Piferrer *et al*., 2009; Zhou and Gui, 2017). Polyploidisation is a chromosome set manipulation and not genetic modification technique and can occur spontaneously in wild fish stocks (Zhou and Gui, 2017). Tetraploidy or triploidy can be induced in many teleost species by inhibiting first cell division or extrusion of the second polar body from zygote, respectively, using a thermal or pressure shock (Maxime, 2008; Piferrer *et al*., 2009). In addition, chemicals such as cytochalasin B can induce triploidy in catarina scallop (*Argopecten ventricosus*) (Ruiz-Verdugo *et al*., 2000). Tetraploid fish can be used as broodstock to produce triploid progeny by breeding with diploid fish, but the low resulting survival rates and the poor viability of gametes make it difficult to use in most teleost species (Yoshikawa *et al*., 2008).

Most triploid fish showed defective gametogenesis and gonadal development (Piferrer *et al*., 2009). The reduced reproductive capacity and sexual maturation caused by triploidisation can help protect wild fish stocks and can result in boosted performances in farmed stocks by suppressing sexual maturation with subsequent positive effects on

somatic growth. Triploidy is commercially exploited in Cyprinidae and Salmonidae where the sterile fish can show better growth and flesh quality (Arai, 2001; Piferrer *et al.*, 2009; Taylor *et al.*, 2013; Zhou and Gui, 2017). GM salmon called AquAdvantage produced by AquaBounty Technologies Inc. (Fortune, PE, Canada) also exploited triploidisation to prevent potential interbreeding of transgenic salmon with wild populations. Triploidisation is feasible from a regulatory standpoint as a non-GM sterilisation method; however, consumer perception and acceptance remain a concern. Triploidy is not feasible in species like tilapia, as there are difficulties in obtaining newly fertilised eggs in large-scale to induce triploidy or maintaining tetraploid broodstock to produce triploid tilapia (Myers, 1986; Penman *et al.*, 1987). It was also shown that triploid tilapia had reduced growth and survival rates when they were treated as same as diploid tilapia (Penman *et al.*, 1987) but there is possibility that the performance of triploid tilapia could be improved by optimisation of rearing condition as suggested in the case of triploid salmon (Benfey, 2016; Fraser et al., 2012; Maxime, 2008).

Gamma irradiation is also able to induce sterility by causing genotoxic and reprotoxic effects (Adam-Guillermin *et al.*, 2012). A high level of sterility was achieved in rainbow trout by exposing the eggs to low levels of gamma radiation (Konno and Tashiro, 1982). The optimal timing of irradiation treatment is important to induce high sterility and low mortality. The sterilisation success of gamma irradiation varies depending on the timing of irradiation during development, as shown in Atlantic salmon, where an increasing sterilisation effect of irradiation was reported at advanced developmental stages (*e.g.* after eyeing stage) than earlier developmental stages, but also the deformity proportion showed a developmental stage-specific sensitivity (*e.g.* high deformity prior to eyeing stage) (Thorpe *et al.*, 1987). In the Amazon molly (*Poecilia formosa*), an acute dose (10 Gy) of irradiation tested in three-month-old fish arrested ovarian maturation without impacting on survival compared to control fish (Woodhead and Setlow, 1980). In addition, chronic exposure (10 days to 30 weeks) to gamma radiation (29 – 178 mGy/day) was shown to induce germ cell losses in several teleost species including Chinook salmon (*Oncorhynchus tshawytscha*), Ameca (*Ameca splendens*), medaka (*Oryzias latipes*), guppy (*Poecilia reticulata*), zebrafish (*Danio rerio*) and European plaice (*Pleuronectes platessa*) (reviewed in Adam-Guillermin *et al.*, 2012). Since this treatment can cause mortality and abnormal development in the irradiated animals, optimisation of sterilisation rate and survival rate is required by adjusting the exposure time, the dose and the developmental stage at which fish are exposed. It is a

relatively simple method to apply which can be optimised with relatively little research effort; however, it is costly and there are public concerns associated with the consumption of irradiated fish.

Surgical gonadectomy can be a very effective sterilisation method when performed optimally and it can be used in most species (Donaldson *et al*., 1993). However, this technique can only be performed in fish of a certain size and it cannot be implemented on a large-scale, only for small scale experiments (Donaldson *et al*., 1993). In addition, it is a very labour-intensive and time-consuming process which may result in fish welfare concerns and mortality (Kapuscinski and Patronski, 2005).

Some hybrid crosses can also cause sterility which acts as a post-zygotic reproductive barrier and is used to maintain speciation (Mallet, 2007; Yoshikawa *et al*., 2018). In some cases, hybrid fish are able to develop gonads, but they often fail to produce viable gametes due to the difference in the meiotic germ cell karyotype and chromosome structure (Hooe *et al*., 1994; Shimizu *et al*., 1997).

Hybrids are used in some tropical fish farming either to improve a given trait (*e.g.* growth, external appearance) or to address reproductive dysfunction in males such as in cachara (*Pseudoplatystoma reticulatum*) which are crossed with ♂ pintado, (*Pseudoplatystoma corruscans*) (reviewed in FAO, 2016). In a recent study, primordial germ cells (PGCs) of a hybrid between blue drum (*Nibea mitsukurii*) and white croaker (*Pennahia argentata*) failed to proliferate in the gonadal anlagen, showing an abnormal nuclear morphology. This hybrid sterility was caused by impaired mitotic cell division of hybrid PGCs, but not by abnormal chromosome pairing during meiosis (Yoshikawa *et al*., 2018). Germ cells may have unique mechanisms to inhibit mitotic cell divisions when the chromosomes contain genetic material of a different species so that proliferative defects of PGCs in hybrid animals could act as a post-zygotic reproductive barrier for maintaining speciation in nature (Yoshikawa *et al*., 2018). In the past, interspecific crosses were used to produce all-male populations in tilapia (Hickling, 1960). All-male tilapia was produced from several interspecific combinations and one of them was *O. niloticus* ✕ *O. aureus* cross, which was used commercially in some fish farms; however, the results were inconsistent, showing mixed sex progeny (Beardmore *et al*., 2001). In most cases, hybridisation often failed to achieve a high proportion of male progeny in mass production due to contamination of the broodstock (Beardmore *et al*., 2001; Conte *et al*., 2017).

Steroid treatment has been tested for sterilisation in the past. In common carp (*Cyprinus carpio*) strains, oral administration of MT (300 – 600 ppm/kg) or testosterone undecanoate (400 ppm/kg) from 1 dah to 30 dah resulted in 74 – 98% sterile fish (Ali and Rao, 1989; Bharadwaj and Sharma, 2000; 2002). However, it is difficult to achieve complete sterility by hormonal treatment and this method is mainly used to manipulate sex either directly or indirectly. The best example of the use of this technique is in tilapia. A strong sex dimorphism in growth can be found in Nile tilapia with males growing up to two-fold faster than females. Therefore, all-male production is preferred in aquaculture. In tilapia hatcheries, MT is widely used for sex reversal to make all-male stock. The in-feed administration (direct sex reversal) of the hormone starts at first feeding for three to four weeks with 60 mg MT/kg of feed, resulting in 95 – 100% masculinisation (Rakocy, 2005). The production of directly sex reversed fish through the use of hormone is banned in Europe and in many major tilapia producing countries for food safety and environmental reasons (Council of the European Union, 2003); however, this practice is very well spread in many tilapia producing countries as it is an efficient and relatively easy method to apply. On the other hand, all-male tilapia can be produced by using YY supermales as broodstock (Beardmore *et al*., 2001). Hormonal treatment-mediated sex reversed XY neo-females can produce YY supermales by crossing with XY males and thereafter can produce GMT (genetically male tilapia) by crossing with XX females (Scott *et al*., 1989). Such indirect sex reversal approach is environmentally friendly and commercially available (Beardmore *et al*., 2001). Indirect sex reversal using hormonal treatment successfully produced monosex progeny in rainbow trout (Bye and Lincoln, 1986), tilapia (all-male) (Scott *et al*., 1989), silver barb (*Puntius gonionotus*) (all-female) (Pongthana *et al*., 1999) and halibut (*Hippoglossus hippoglossus*) (all-female) (Palaiokostas *et al*., 2013a).

Exposure to high temperature can also suppress reproductive development and induce sterility in some cases when applied at different developmental stages from larvae to adults in teleosts species including Argentinian silverside (*Odontesthes bonariensis*) and Patagonian freshwater pejerrey (*Patagonina hatcheri*), Nile tilapia, Mozambique tilapia (*O. mossambicus*), yellowtail tetra (*Astyanax altiparanae*) (de Siqueira-Silva *et al*., 2015; Ito *et al*., 2003; Nakamura *et al*., 2015; Pandit *et al*., 2015; Strüssmann *et al*., 1998). Female Nile tilapia exposed to an incipient sublethal high temperature of 37°C for 45 - 60 days appear to become sterile and are devoid of germ cells at all stages of the life cycle (Pandit *et al*., 2015). In addition, male Mozambique tilapia exposed to 37°C for 50

days during the sexual determination and differentiation window were infertile with testes lacking germ cells as adults (Nakamura *et al*., 2015). This method, however, is not commercially applicable as a means to induce sterility as it can result in high mortalities, impact on the overall welfare of the treated stocks and is not fully reliable.

Chemical treatment can also be used as a means of sterilisation in fish. A combined treatment of heat with chemical treatment such as cytotoxic busulfan can effectively induce sterility even in adulthood as well as larval stage in fish including Nile tilapia, Patagonian pejerrey (*Odontesthes hatcheri*), common carp and zebrafish (Lacerda *et al*., 2006; Majhi *et al*., 2009, 2017; Nóbrega *et al*., 2010). This combined treatment appears to be effective at sterilising stocks and it is widely used to produce sterile recipient fish for germ cell transplantation studies (Golpour *et al*., 2016; Lacerda *et al*., 2006; Majhi *et al*., 2009); however, it is not suitable for application in aquaculture for food safety reasons.

Other chemicals which interfere with the normal endocrine control of reproduction either through an inhibition or an over-stimulation can also result in sterility. One such example is by targeting the Gonadotropin releasing Hormone (GnRH) produced by the hypothalamus (patent US 7,194,978 B2) (Zohar *et al*., 2007). This method consists in disrupting the GnRH system by administering GABA (γ-Aminobutyric acid), GABA receptor agonist, or GABA receptor antagonist during early development in fish (Zohar *et al*., 2007). However, the commercial application of this technique is very limited due to the potential side effects of the treatments on other developmental processes and physiological functions which have not yet been characterised. In addition, the cost of the treatment is prohibitive in most cases. Another concept is to develop and administer vaccines against proteins controlling the survival of PGCs or gametes. This immunisation approach has been the object of intensive research in a recently funded project by the Norwegian Research Council "SalmoSterile". This targeted autoimmunity strategy against gonadal proteins was tested in zebrafish with effects reported on the development of gonads (Presslauer *et al*., 2014). However, reports so far are inconsistent and treatments are very costly which prevents such a technique from being commercially implemented.

While these non-GM sterilisation methods including chromosome set manipulation, gamma irradiation, steroid or chemical treatment are relatively easy to test and require relatively less research efforts than traditional transgenic fish, in most cases, they are not 100% effective at inducing sterility; therefore, requiring individual screening. In addition,

many of the above can lead to welfare concerns, abnormal development and increased mortalities while also raising food safety concerns. Triploidy induced by pressure shocking in salmonids may be an exception due to extensive research efforts over the last decades; however, large-scale implementation within existing selective breeding programs remains problematic and it is only available for a limited number of farmed species due to the need to control fertilisation through stripping of gametes.

1.1.2 Genetically modified (GM) sterilisation methods

The traditional way to make transgenic fish is costly to develop and takes a long time to establish the transgenic line with desirable traits. In addition, the resultant GM transgenic animals contain foreign genes which are originated from different organisms. This type of GM animals poses public concerns; therefore, the approval for human consumption is hard to obtain. Today, the only example of a sterile GM animal that recently obtained approval from the US Food and Drug Administration (FDA) and the Canadian Food Inspection Agency for commercialisation and consumption is the AquAdvantage salmon originally developed in 1989 by AquaBounty Technologies Inc. (Ledford, 2015). The genetically engineered salmon can grow faster than a conventional farmed salmon due to the inclusions of a growth hormone gene from Chinook salmon, the expression of which is regulated by an antifreeze gene promoter from the Ocean pout (*Zoarces americanus*).

One of the transgenic sterilisation methods is the maternal sterilisation technology developed and patented by Lauth and Buchanan (2012). They developed a maternal sterility construct (MSC) consisting of three components: (1) ovarian germ cell-specific gene promoters, such as *zona pellucida* (*zpc*) or *askopos* (*kop*) promoters, (2) open reading frame (ORF) of proapoptotic gene, *bax* (*B-cell lymphoma 2 associated X*) and (3) 3' untranslated region (UTR) of *dnd* (*dead end*). Thus, the expression of the *bax* gene would be limited in the germline, inducing apoptosis in germ cells. Hemizygous MSC transgenic male fish can be crossed with wild type (WT) females to produce hemizygous MSC transgenic females, which are fertile lineage ending females. This lineage ending female can produce infertile progeny by crossing with WT or hemizygous MSC transgenic male. The potential problems of this method were reviewed by Wong and Zohar (2015a) including (1) unintended apoptosis in oocytes or somatic cells which can reduce the fecundity of lineage ending female or induce high mortality, respectively, and (2) time and labour intensive maintenance of the broodstock as it requires production of

both hemizygous MSC transgenic males and females and a screening procedure for lineage ending females.

In recent years, inducible transgenic sterilisation methods by targeting primordial germ cells (PGCs) were reported in fish. Wong and Collodi (2013) developed and patented the transgenic fish line in which sterility can be induced by inhibiting the migration pathway of PGCs. They developed a transgenic line of zebrafish, that expresses C-X-C motif ligand 12a (*cxcl12a*) under the control of the heat-shock protein 70 (*hsp70*) promoter and 3'UTR of *nanos3* (RNA binding zinc finger protein). Heat treatment at the 4- to 8-cell stage at 34.5°C for 18 hours causes disruption of endogenous Cxcl12a signal, which guides PGCs for successful migration, resulting in infertile males with under-developed gonads, lacking gametes in adulthood. This approach allows the maintenance of a fertile broodstock as well as the production of sterile fish. While this technique is promising for implementation at a commercial scale (Wong and Collodi, 2013), the induction of mis-migration of PGCs by disrupting *cxcl12a* gradient signal has potential risks. The resultant ectopic PGCs which failed to migrate into the genital ridge undergo apoptosis, which is important to prevent the formation of germ cell tumours as reported in mammals (Göbel *et al*., 2000; Schneider *et al*., 2001; Ueno *et al*., 2004). However, the ectopic PGCs that survived may pose a potential health risk. Furthermore, the Cxcl12a/C-X-C motif receptor 4b (Cxcr4b) pathway is likely to be involved in other important developmental processes. This was shown in mice where Cxcr4 or Cxcl12a knockout (KO) mice revealed that these genes also play a critical role in heart and brain development, hematopoiesis and vascularisation (reviewed in Ratajczak *et al*., 2006). In addition, the induction temperature requires species-specific optimisation and must be kept strictly constant otherwise, it could lead to reduced sterilisation success (Wong and Zohar, 2015a).

Exploitation of nitroreductase/prodrugs system has been tested as a means of transgenic sterilisation. Nitroreductases can convert non-toxic prodrugs into cytotoxic metabolites, inducing programmed cell death which has been used in cancer therapy (Searle *et al*., 2004). Spatial, temporal and germ cell-specific induction of cell elimination can be achieved by employing gonad specific promoters to drive nitroreductase expression so that the prodrugs are only converted into cytotoxic metabolites in nitroreductase-expressing germ cells (Curado *et al*., 2008; Hsu *et al*., 2010; Hu *et al*., 2010; Zhou *et al*., 2018). However, 100% sterility is hard to achieve with this technique and when the promoter is testis- or ovary-specific, only one sex can be sterilised, which

can prevent large-scale applications. In addition, copper transport protein gene promoter which is activated by the presence of copper (Li *et al*., 2018; Su *et al*., 2015) or salt sensitive repressible promoters (Li *et al*., 2017c) were tested to control the expression of short hairpin RNA (shRNA) to knockdown germline specific genes such as *nanos* and *dnd* in channel catfish (*Ictalurus punctatus*) to induce sterility. However, when the transgenes are controlled via a repressor or an inducer such as temperature treatment, prodrugs, copper or salt to eliminate PGCs (Hsu *et al*., 2010; Hu *et al*., 2010; Li *et al*., 2017c; Li *et al*., 2018; Su *et al*., 2015; Zhou *et al*., 2018), there are potential risks that the unexpected presence of repressor or inducer molecules in natural waters may affect the expression of the transgene, which can then cause unpredictable results.

There have been various sterilisation methods in fish with or without genetic modification. Transgenic fish may not be as socially acceptable as non-transgenic sterilisation methods. However, the recent access to simple and efficient genome editing tools, such as ZFNs and TALENs and clustered regularly interspaced short palindromic repeats/CRISPR associated effector 9 (CRISPR/Cas9) system, means it is possible to produce genome edited animals/lines over a relatively short time with desired commercial traits without introducing transgenes or for gene functional analysis studies in fish (Zhu and Ge, 2018). These are thought to have great potential in aquaculture. The genetically engineered animals produced by these gene editing technologies do not have any foreign genes (transgenes) in most cases and the changes can be as small as for natural mutation or selective breeding, and generally smaller compared to the genetic changes that occur in traditional GM organisms (GMO). Currently, they are considered as GMO in EU; however, an appropriate new regulation will be required for organisms which have only a few changes in their genome produced by targeted genome editing (Eriksson *et al*., 2018).

Reproductively sterile farmed fish have clear advantages in production including fish quality and better growth performances, prevention of cross-breeding between escapees and wild stocks and intellectual property (IP) protection for pedigrees from selective breeding programs. The reproductive containment is not only important for farmed species but also for the prevention of unauthorised animals from colonising an environment such as non-native, invasive or patented species. In addition, sterility is a requirement for research on germ cell transplantations as recipients, which aim to conserve the genetic stocks of endangered or valuable species (Golpour *et al*., 2016).

1.2 Primordial germ cells (PGCs) and gonadal development

1.2.1 Definition and importance of PGCs

Germ cells are a unique cell type that belongs to the germline, which enables the generation of offspring and inheritance of genetic information by producing gametes (Extavour and Akam, 2003). The germline precursor undergoes asymmetric cell division to give birth to primordial germ cells (PGCs) (Knaut *et al.*, 2000). At this stage, the aggregates are referred to as presumptive PGCs (pPGCs) and they are not morphologically distinguishable from soma lineage until late blastula (Raz, 2003). As a result of the asymmetric cell division of pPGCs, two different daughter cells are formed; one is committed to soma and the other to germline (Xu *et al.*, 2010a). Only when both cells are committed to the germline by symmetrical cell divisions, is the parental cell called a PGC (Knaut *et al.*, 2000; Xu *et al.*, 2010a). PGCs are set aside from somatic cells and migrate into prospective gonadal sites during the early developmental period (Raz, 2003). After they coalesce with somatic cells into the gonad, PGCs give rise to the germ cell lineage and finally become gametes through sexual differentiation and gametogenesis (Cinalli *et al.*, 2008). After fertilisation, the zygote repeats this process which enables reproduction (Cinalli *et al.*, 2008).

1.2.2 PGCs specification

The specification of germ cells in animals includes two different modes (reviewed in Extavour and Akam, 2003; Whittle and Extavour, 2017). Firstly, the germ cell can be specified by the maternally provided determinants in germ granules and this mode of specification is called 'inheritance' and can be found in germ cells from flies, worms, fish and frogs (Lehmann, 2012; Whittle and Extavour, 2017). Germ plasm (GP) is evolutionarily conserved, non-membranous cytoplasmic structures containing maternally provided ribonucleoproteins (RNPs) which are involved in germline specification in *Drosophila*, *Caenorhabditis elegans*, *Xenopus* and teleosts, while these are not involved in germline specification in mammals as GP appear later during differentiation stages of mammalian germ cells (reviewed in de Mateo and Sassone-Corsi, 2014). This mode of specification called 'induction', is found in mammals and most species of animal phyla, is independent of maternal information and the specification

process is affected by inductive signals from surrounding tissues such as bone morphogenetic protein (BMP) signalling (Hogan, 1996) and canonical WNT (wingless-type MMTV integration site family, member 4)/β-catenin signalling pathways (Lehmann, 2012; Saitou and Yamaji, 2012; Strome and Updike, 2015; Whittle and Extavour, 2017).

In fish, the mode of PGC specification is inheritance with the maternally deposited determinants (reviewed in Yamaha *et al*., 2010). The maternal transcripts are located in peripheral aggregates (germ granules) in the blastodisc in zebrafish (Theusch *et al*., 2006) (Fig. 1.1). These maternal mRNA aggregates are associated with actin filaments and they move to the end of the first cleavage furrow by microtubule-dependent movements during cytoskeletal rearrangement (Theusch *et al*., 2006). Yamaha *et al*. (2010) reported that an inconsistency on the localisation of *vasa* mRNA-containing aggregates in early cleavage stages between ice goby (*Leucopsarion petersii*) (Miyake *et al*., 2006) and goldfish (*Carassius auratus*) (Otani *et al*., 2002), suggesting that depending on the distribution of cytoplasm/yolk and the patterns of cleavage, the localisation and timing of the maternal aggregates which determine PGCs can vary in teleosts (Yamaha *et al*., 2010).
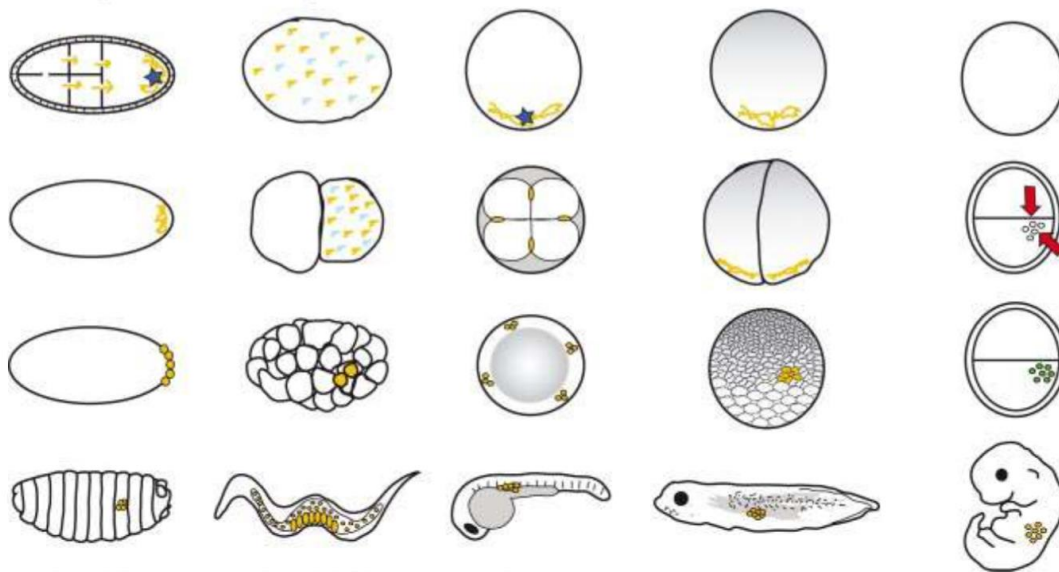
Figure 1.1. Schematic diagram of germ cell specification and the localisation during ontogenic development. From the top to bottom is the chronological order of development. In mice, somatic signals (arrows) induce the formation of PGCs, while in the rest of organisms, PGCs are specified by GP components (threads or dots). Dark blue star indicates Oskar in *Drosophila* and Bucky ball in zebrafish, which autonomously assemble GP components. Figure from Ewen-Campen *et al*. (2010).

To acquire germ cell-specific functions and totipotency, germ cells inhibit somatic differentiation and take advantage of germline-specific RNA regulation (Cinalli *et al*., 2008). Despite different modes of PGC specification, there are common conserved molecular mechanisms to maintain PGC fate such as differential transcriptional silencing between PGCs and soma cells for PGCs-specific gene expression (reviewed by Richardson and Lehmann, 2010). In zebrafish, *nanos3* is maternally deposited in GP and plays an important role in specification and maintenance of PGCs (Knaut *et al*., 2000; Köprunner *et al*., 2001; Yoon *et al*., 1997). One of the molecules which are also abundant in the GP aggregates (nuage) is piwi-interacting RNAs (piRNAs)-related proteins, Piwi (P-element induced wimpy testis), which protect genome integrity by inhibiting the activity of transposable elements via germline-specific RNA silencing (Marlow, 2015). More evidence supported that post-transcriptional gene regulation is particularly important in the germline, which includes piRNA-mediated transposon silencing, ribonucleoprotein formation to repress translation and chromatin architecture-mediated gene regulation (Cinalli *et al*., 2008; Ewen-Campen *et al*., 2010). Even after the specification, PGCs require specific mechanisms to maintain the germline fate, otherwise, they can revert to somatic cells (Strome and Updike, 2015). However, the molecular network of the germline is still not well-defined yet and remains to be elucidated.

1.2.3 PGCs migration

1.2.3.1 Initiation of PGCs migration

Following specification, PGCs acquire motility and receive directional signals to begin migration through multiple steps (Blaser *et al*., 2005). In zebrafish, PGCs show a simple round shape at 3 hours post fertilisation (hpf) (Fig. 1.2) (Blaser *et al*., 2005). Then, PGCs display complex non-polarised morphology with random protrusions at around 3.5 hpf (Blaser *et al*., 2005). At the third step, they finally become responsive to chemokine Cxcl12a, showing a polarized shape with one protrusion at 4.5 hpf (Blaser *et al*., 2005). The transition to the third step requires several molecular changes. At first, *de novo* zygotic transcription in PGCs is necessary, since cells treated with α-amanitin, an inhibitor of RNA polymerase II and III, failed to enter the third stage (Blaser *et al*., 2005). This step also requires translation of *dnd*, an RNA-binding protein, as knockdown of *dnd* hindered the transition (Blaser *et al*., 2005; Weidinger *et al*., 2003). At this stage, E-

cadherin level is down-regulated, and the removal of *dnd* in PGCs results in no change in the level of E-cadherin and non-motility of PGCs, indicating the involvement of *dnd* in the down-regulation of E-cadherin (Weidinger *et al*., 2003). However, how *dnd* regulates E-cadherin remains to be elucidated.
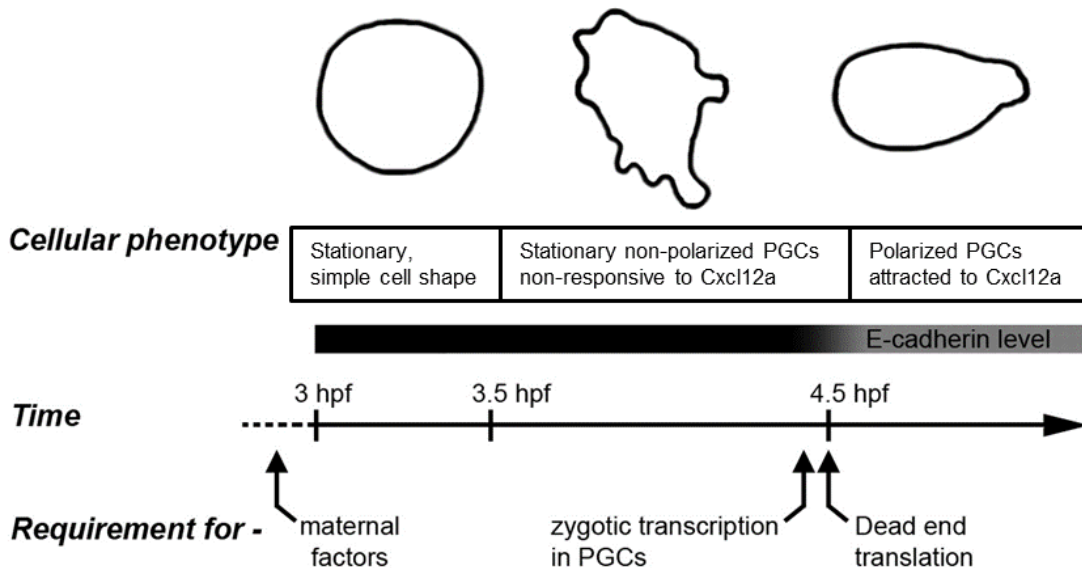
Figure 1.2. Initiation of PGCs migration in zebrafish. Figure from Blaser *et al*. (2005).

1.2.3.2 Pathway of PGCs migration

Weidinger *et al*. (1999) described the migration pathway to the gonad of zebrafish PGCs by six distinct steps (6 - 24 hpf): (1) migration to the dorsal side of the embryo (Fig. 1.3, 4.5 hpf), (2) exclusion from the dorsal midline, (3) alignment with the anterior and lateral mesoderm, (4) formation of two lateral PGC clusters at somites 1-3 (Fig. 1.3, 10.5 hpf), (5) anterior migration of trailing PGCs to join the main PGC clusters and (6) posterior positioning of PGC clusters at somite 8 (Fig. 1.3, 13 hpf).
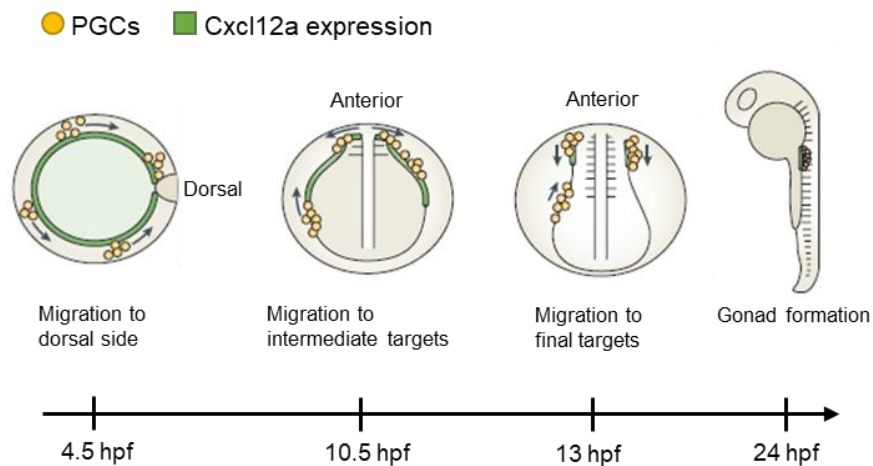
Figure 1.3. Stages of primordial germ cell migration in zebrafish. After specification, PGCs migrate dorsally (animal pole view). At gastrulation, 4.5 hpf, PGCs follow the expression of Cxcl12a and at 10.5 hpf PGCs reach somites 1-3 (intermediate targets, lateral view, left side). PGCs migrate towards somites 8-10 at 13 hpf (frontal view). At 24 hpf, PGCs form gonadal anlagen (lateral view, left side). Figure from Richardson and Lehmann (2010).

The migration of PGCs is guided by chemokine Cxcl12a and its cognate receptors, Cxcr4b expressed in PGCs and Cxcr7b which is ubiquitously expressed in the embryo and mainly functions in somatic cells (Boldajipour *et al*., 2008; Doitsidou *et al*., 2002; Knaut *et al*., 2003; Mahabaleshwar *et al*., 2008) (Fig. 1.4). Cxcl12a (previously named stromal cell derived factor 1a, Sdf1a) is produced by neighbour soma cells (Doitsidou *et al*., 2002; Knaut *et al*., 2003). Both Cxcr4b and Cxcr7b belong to seven transmembrane G protein coupled receptor (GPCR). The gradient of Cxcl12a signal is crucial for the proper guidance for the migration of PGCs, and Cxcr7b establishes the signal gradient. Cxcr7b located in the intracellular structure of soma cells gets rid of Cxcl12a signal from the somatic tissues by coupling with them, providing a gradient signal pattern of Cxcl12a to Cxcr4b located in the membrane of PGCs (Boldajipour *et al*., 2008) (Fig. 1.4). In response to Cxcl12a signal, Cxcr4b located in zebrafish PGCs activates polarizing, and it leads to an increase in local free calcium. Subsequently, it activates myosin contraction that generates bleb-like protrusions by causing cytoplasmic flow, and these blebs show the direction of the cell migration (Blaser *et al*., 2006). This is different from many other

migratory cells where actin polymerisation is involved in the formation of their cellular protrusions (Richardson and Lehmann, 2010).

Cxcl12a/Cxcr4 signalling in zebrafish plays an important role in various processes as well as the migration of PGCs. The Cxcl12a/Cxcr4/ is required for the migration of sensory cells of the lateral line (David *et al*., 2002; Li *et al*., 2004), the guidance of retinal ganglion cell axons within the retina (Li *et al*., 2005), placode assembly and sensory axon pathfinding (Miyasaka *et al*., 2007), and migration of GnRH3 neurons in zebrafish (Palevitch *et al*., 2010). Knockdown of Cxcl12a or its receptor Cxcr4 results in random migration of PGCs (Doitsidou *et al*., 2002). This migration signal pathway of PGCs was exploited by Wong and Collodi (2013) to produce large numbers of infertile fish by inducing ubiquitous expression of Cxcl12a in zebrafish embryos resulting in the disruption of the normal PGC migration pattern.
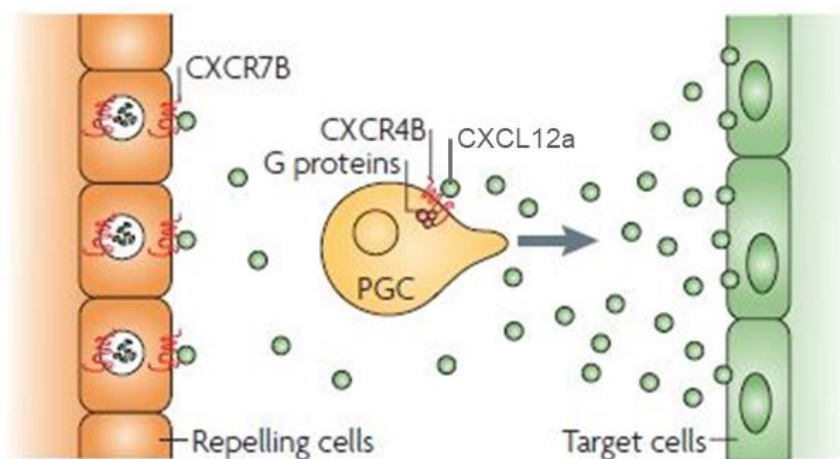
Figure 1.4. Molecular regulation of PGC migration in zebrafish. PGCs are expressing Cxcr4b and migrate towards the Cxcr4b ligand, Cxcl12a, secreted by somatic cells. Cxcr7b promotes the degradation of Cxcl12a, generating proper gradient signal for the migration of PGCs. Figure adapted from Richardson and Lehmann (2010).

1.2.4 PGCs and sex determination and differentiation

Once PGCs have completed their migration and reached the gonadal anlagen, they are regulated by both intrinsic and extrinsic (environmental) factors and proliferate prior to sex determination (Campolo *et al*., 2013; Chassot *et al*., 2017). In contrast to mouse PGCs, fish PGCs do not proliferate during migration, but they become mitotically active

in the gonadal anlagen and increase in number prior to sex differentiation (Braat *et al*., 1999; Yoshikawa *et al*., 2018).

It is known that the interaction between germ cells and gonadal somatic cells plays an important role in the early gonadal formation and sex differentiation in vertebrates (Goto *et al*., 2012). In mice, their sex is determined by the surrounding environmental factors: RNA helicase MVH (mouse Vasa homolog) and WNT-4 signalling are required for PGC proliferation in XY and XX gonads, respectively (Tanaka *et al*., 2000; Vainio *et al*., 1999). PGCs become gonocytes by 12.5 days post coitum (dpc) in XY gonads or become oogonia and enter meiosis at around 13.5 dpc in XX gonads under the control of the surrounding environment (Adams and McLaren, 2002; Chassot *et al*., 2017). In teleosts, *gsdf* (*gonadal soma-derived factor*) is essential for testicular differentiation, which is downstream of *dmy* (DM-domain gene) in medaka or *dmrt1*(*double sex- and mab3-related transcription factor 1*) in tilapia in the sex-determining cascade (Jiang *et al*., 2016; Myosho *et al*., 2012). On the other hand, the lack of male sex determination gene on X chromosomes triggers female sex determination and differentiation and promote expression of female fate-related genes, including *cyp19a1* (*cytochrome P450 family 19 subfamily A member 1*), *foxl2* (*forkhead box l2*), *sf-1* (*steroidogenic factor-1*) and *wnt4* (Mei and Gui, 2015).

nanos2 and *nanos3* (RNA binding zinc finger proteins) or *dnd* KO mice showed loss of germ cells, but normal testicular soma differentiation implying that germ cells are not essential for the primary sex determination or differentiation in mammals (Goto *et al*., 2012; Tsuda *et al*., 2003; Youngren *et al*., 2005). Likewise, in some teleosts, the absence of germ cells did not affect the gonadal soma differentiation. In goldfish and loach (*Misgurnus anguillicaudatus*) the loss of germ cells by knockdown of *dnd* during embryogenesis showed that absence of germ cells did not affect the development of sexually dimorphic gonads, indicating that germ cells are not the primary factor for sex determination in these species (Fujimoto *et al*., 2010; Goto *et al*., 2012). On the other hand, germ cell deficient zebrafish induced by knockdown of *dnd* all developed as males with normal expression of testicular genes such as *amh* (*anti-Müllerian hormone*), *sox9a* (*sex determining region Y-box 9a*) and *11β-HSD* (*11β-Hydroxysteroid dehydrogenase*), demonstrating that undifferentiated gonads lacking germ cells favour phenotypic male sex in zebrafish (Siegfried and Nüsslein-Volhard, 2008; Weidinger *et al*., 2003). This phenomenon was also reported in medaka and Nile tilapia, where germ cell deficient XX fish went through female-to-male sex reversal (Kurokawa *et al*., 2007; Li *et al*., 2014),

which indicates that germ cells are essential for the female fate of gonadal somatic cells in these species. These results indicate that the mode of controlling sex determination and differentiation with or without the interaction between germ cells and somatic cells in teleosts is species-specific. These results also showed that gonadal somatic cells are able to develop as testicular or ovarian tissues expressing gonadal somatic cell-specific genes to some extent without germ cells in sterile fish.

Nile tilapia is a gonochoristic fish which has a complex sex determining system rather than a simple male heterogametic (XX-XY) system in which sex-determining loci are present in linkage groups (LG)1, LG3 and LG23 (Cnaani *et al*., 2008; Khan *et al*., 2015; Palaiokostas *et al*., 2013b). In Nile tilapia, PGCs reach the gonadal anlagen at 3 days after hatching (dah) and the active proliferation was recorded after 9 dah and 14 dah in XX female and XY male, respectively (Kobayashi *et al*., 2002, 2000; Nakamura *et al*., 1998). At 5 dah, the expression of *gsdf* starts to be sexually dimorphic and overexpression of *gsdf* in XX fish induces female-to-male sex reversal (Jiang *et al*., 2016). Sexual differentiation is apparent from 20 and 25 dah in XX and XY gonads, respectively, as the ovarian cavity and the intratesticular efferent duct are formed (Fig. 1.5) (Kobayashi *et al*., 2013). Finally, oogenesis and spermatogenesis initiate between 25 and 30 dah in XX female and between 50 and 70 dah in XY male Nile tilapia (Fig. 1.5) (Kobayashi *et al*., 2000, 2013). It was shown that the presence of PGCs is essential for the female fate of gonadal somatic cells, but not for male fate in this species (Li *et al*., 2014).

In mammals, after the sex determination and differentiation, gonocytes in XY gonads are arrested in G0/G1 (Gap 0/Gap 1 of the cell cycle) and remain quiescent until *sry* (*sex determining region Y*) is expressed to initiate somatic cell differentiation in male gonads (Sekido and Lovell-Badge, 2008). *sry* activates the transcription factor *sox9* and then induces the formation of Sertoli cells from somatic cells (Chaboissier *et al*., 2004). Then, spermatogonial stem cells (SSCs) appear to reside in a unique microenvironment, 'niche', which is provided by somatic support cells in the testis, including Sertoli cells, peritubular myoid cells, Leydig cells and surrounding vasculature (Oatley and Brinster, 2012; Xie, 2008; Yoshida *et al*., 2007). Establishment of appropriate germ cell niches is essential for the development and differentiation from gonocyte into spermatogonium and it was reported in human that the failure can cause a testicular germ cell cancer in adulthood (Skakkebaek, 1972; Rajpert-De Meyts, 2006).
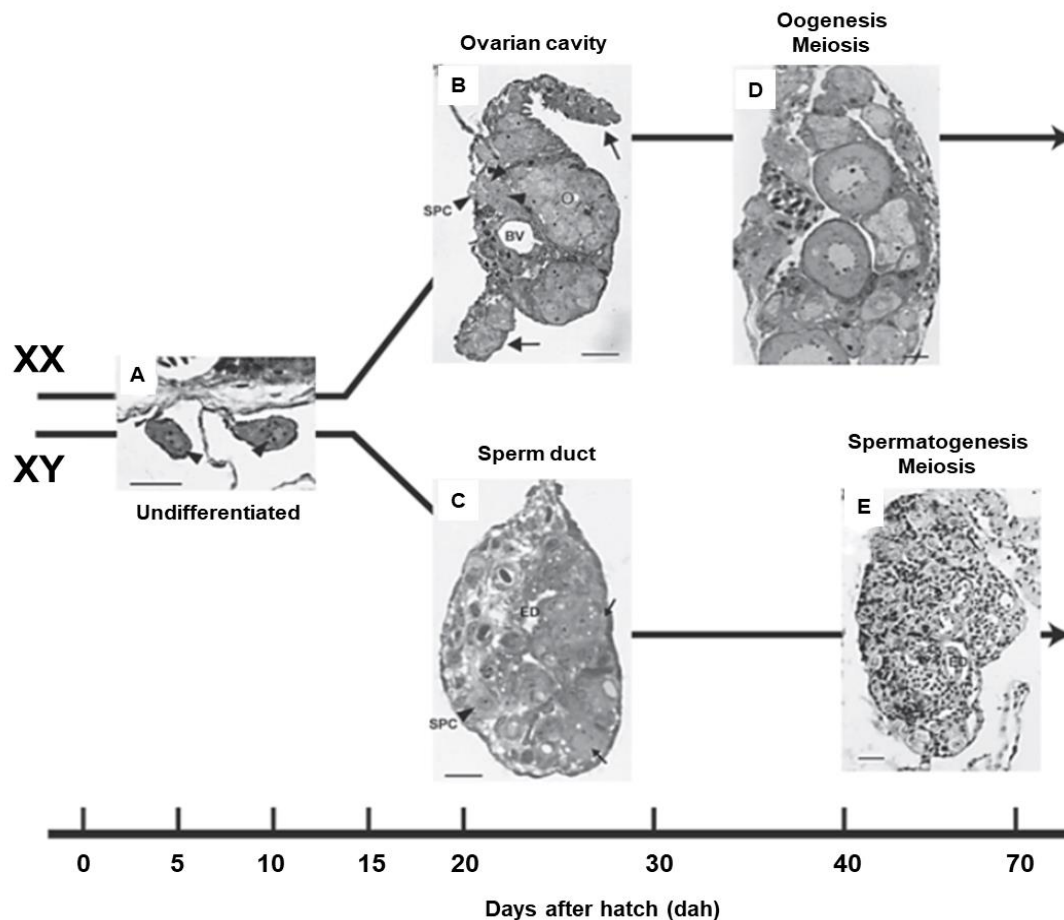
Figure 1.5. Overview of gonadal sex differentiation process in Nile tilapia. (A) Gonad at 3 dah. Arrowheads indicate undifferentiated germ cells. (B) XX gonad at 20 dah. Arrows indicate the elongation of 2 somatic ridges for the formation of the ovarian cavity and arrowheads indicate clusters of steroid-producing cells (SPC) near a blood vessel (BV). (C) XY gonad at 25 dah. ED, efferent duct. Arrows indicate gonial stage of germ cells. (D) XX gonad at 35 dah. (E) XY gonad at 70 dah. Scale bars = 20 µm, except for B (30 µm). Figure from Kobayashi *et al.* (2013).

1.3 Characteristics of PGCs in teleosts

PGCs can be distinguished from soma cells using PGC-specific morphological, biochemical and molecular characteristics during early developmental stages.

1.3.1 Morphological and biochemical features of PGCs

The first identification of piscine PGCs was reported in mummichog (*Fundulus heteroclitus*) embryos in 1921, from peripheral endoderm in embryos at 46 hpf, showing a close relationship with the periblast layer (Richards and Thompson, 1921 cited in Deniz Koç and Yüce, 2012). Since then, PGCs were reported in various species including largemouth bass (*Micropterus salmoides*), at the bulging end of caudal periblast, eastern mosquitofish (*Gambusia holbrooki*), in the caudal area of the blastodisc during early gastrulation and common carp in the mesodermal layer of embryonic disk (reviewed by Deniz Koç and Yüce, 2012).

PGCs have morphologically distinguishable features such as the large size of the cell and the nucleus, the round shape and the slightly stained cytoplasm shown in medaka (Satoh and Egami, 1972). PGCs of dusky grouper, *Epinephelus marginatus,* are large in size (12 - 24 μm in diameter), oval in shape, unstained and often arranged in clusters (Mandich *et al*., 2002). The oval-shaped PGCs with large nucleus were also reported in grey mullet (*Mugil cephalus*) (Chang *et al*., 1995) and ukigori, *Gymnogobius* species with PGC size of 10 - 20 μm (Saito *et al*., 2004). Overall, the common morphological traits of PGCs in teleosts are large round-shaped cells with large nucleus (Deniz Koç and Yüce, 2012). In addition, they commonly have nuage materials which contain electron-dense bodies in their cytoplasm, which were used across phyla as a marker of PGCs at early developmental stages (Herpin *et al*., 2007). In zebrafish, the granules known as nuage materials were stained deeply with PAS and Best Carmin around the nucleus of PGCs (Deniz Koç and Yüce, 2012). These granules around the nucleus were also reported in medaka and rosy barb (*Barbus conchonius*) (Gevers *et al*., 1992). In common carp, PGCs showed a close relationship between nuage material and mitochondria (van Winkoop *et al*., 1992), also reported in pacu (*Piaractus mesopotamicus*) (Abdalla and Cruz-Landim, 2004). These germ cell specific organelles have also been called electron-dense structures, basophilic bodies, mitochondrial clouds, chromatoid bodies, yolk nuclei or Balbiani bodies as well as GP (Extavour and Akam, 2003). In teleosts, PGCs are specified by the maternally provided determinant in GP (Extavour and Akam, 2003). GP is evolutionarily conserved, non-membranous cytoplasmic structures containing RNPs which are involved in germline specification as shown in *Drosophila*, *C. elegans*, *Xenopus* and other teleosts (de Mateo and Sassone-Corsi, 2014; Kloc *et al*., 2014; Tada *et al*., 2012).

1.3.2 Molecular features of PGCs

GP consists of microscopically distinct maternally deposited RNAs and RNA-binding proteins, which are required to maintain germ cell fate and totipotency, otherwise, they become somatic cells (Cinalli *et al*., 2008; Tada *et al*., 2012; Lai and King, 2013). Knockdown of specific RNA molecules of GP including *nanos*, *dnd* and *vasa* genes can lead to severe defects in PGC development, survival and /or migration in fish (Köprunner *et al*., 2001; Li *et al*., 2006; 2009; Weidinger *et al*., 2003).

- vasa

vasa, an ATP-dependent RNA helicase of the DEAD-box family, is one of the well-known components of GP and the first reported molecular marker of PGCs in fish (Hay *et al*., 1988; Yoon *et al*., 1997). In zebrafish, *vasa* mRNAs are synthesized during oogenesis and enriched at the cleavage furrows (planes) during the first two cleavages, and the resultant *vasa*-expressing 4 lines along the cleavage plane at 4-cell stage become PGCs (Yoon *et al*., 1997). Electron microscopic observation also supported that *vasa* mRNAs were fused into electron-dense nuage materials at 4-cell stage showing that GP is segregated at this early stage in zebrafish (Knaut *et al*., 2000). On the other hand, *vasa* expression was uniformly distributed until late gastrulation in medaka, showing a discrepancy in the timing of GP segregation, but the inheritance of maternally provided asymmetrically-localized cytoplasmic determinants directs cells to assume the germline fate similar to zebrafish PGCs (Herpin *et al*., 2007). *vasa* was the first germline-specific gene investigated in Nile tilapia, being reported to play an important role in the formation and migration of PGCs (Kobayashi *et al*., 2000). *vasa* mRNAs showed different expression pattern depending on the sex and developmental stage of germ cells in Nile tilapia: in testis, *vasa* signals are strong in spermatogonia and decrease in early primary spermatocytes, followed by no evident expression in the advanced stages of spermatogenesis, while *vasa* is expressed strongly in oogonia to previtellogenic oocytes, followed by a decrease from vitellogenic to postvitellogenic oocytes in the ovary (Kobayashi *et al*., 2000). Although the expression level of *vasa* reduced from vitellogenic stages, *vasa* transcripts were still detectable in full-grown oocytes and ovulated eggs, suggesting that *vasa* is maternally deposited.

- dnd

Dnd is an RNA-binding protein, which protects germ cell-specific RNAs from microRNA mediated degradation by binding to the 3'UTRs of germ cell-specific genes in human cell and zebrafish PGCs (Kedde *et al*., 2007). In Atlantic cod, knockdown of *dnd* caused a decrease in expression levels of *vasa*, *nanos3* and *tdrd7* (*tudor domain containing protein 7*) (Škugor *et al*., 2014b). Furthermore, in zebrafish, *dnd* knockdown resulted in transdifferentiation of PGCs into somatic cells, showing that *dnd* plays an essential role in the maintenance of germ cell fate (Gross-Thebing *et al*., 2017). While *dnd* knockdown did not result in complete loss of germ cells in Atlantic cod and zebrafish (Škugor *et al*., 2014b; Gross-Thebing *et al*., 2017), *dnd* KO using CRISPR/Cas9 resulted in germ cell loss in Atlantic salmon (Wargelius *et al*., 2016), implying the importance of zygotic transcripts of *dnd* in survival of germ cells. The Dnd protein protects the transcripts of PGC-specific genes from the inhibitory function of microRNAs (Kedde *et al*., 2007). In zebrafish, miR-430 is expressed in both somatic and germ cells from the onset of zygotic transcription and down-regulates the maternal mRNAs such as *nanos* and *tdrd7* in somatic cells, but its binding and degradation are inhibited in germ cells by Dnd which binds to target 3'UTRs (Giraldez *et al*., 2005, 2006; Mishima *et al*., 2006). Their 3'UTR is evolutionary conserved in teleosts and contains regulatory elements involved with RNA localisation, stability and/or translation (Knaut *et al*., 2002; Škugor *et al*., 2014a). In Nile tilapia, automated computational analysis by NCBI revealed predicted sequences for one *dnd1*, but their molecular characteristics remain unknown in this species.

- dazl and boule

daz (*Deleted in Azoospermia*) and *dazl* (*daz-like*) genes are required for the development of PGCs and differentiation and maturation of germ cells, and the expression of the *daz* gene family is restricted to germ cells (reviewed by Yen, 2004). In teleosts, *dazl* homologs were found in medaka (Xu *et al*., 2007), zebrafish (Maegawa *et al*., 1999), gibel carp (*Carassius auratus gibelio*) (Peng *et al*., 2009), rainbow trout (Li *et al*., 2011) and half-smooth tongue sole (*Cynoglossus semilaevis*) (Wang *et al*., 2014a). In Nile tilapia, two daz family genes, *boule* and *dazl*, were identified and showed germline specific expression in both sexes (Bhat and Hong, 2014). They showed stage-specific expression: in ovary, *boule* and *dazl* were expressed from stage I to stage III with peak expression taking place at stage III and II, respectively. In testis, *boule* was detectable in spermatogonia, spermatocytes and spermatids, while *dazl* signal was weak in

spermatogonia but abundant in spermatocytes, implying they could be used as potential markers to stage of gametogenesis in adult tilapia (Bhat and Hong, 2014). Their expression pattern in PGCs during early developmental stages and their function in tilapia PGCs remain to be elucidated.

- nanos

nanos are well conserved germ cell markers across phyla together with *vasa*. *nanos* genes were predicted in Nile tilapia and their sequences were compared with other species showing highly conserved sequence within zinc finger domains located in the C-terminal domain (Bellaiche *et al*., 2014). The canonical Nanos protein contains two CCHC zinc-finger domains which bind to RNA with low sequence specificity. Puf (Pumilio and FBF, fem-3 mRNA-binding factor) domain of the RNA binding protein Pumilio provides the specificity to Nanos by binding specific target sequences through interacting with Nanos protein, and the Nanos/Pumilio complex represses translation of target RNAs by deadenylation (Ewen-Campen *et al*., 2010). The different roles were reported depending on species. In *Drosophila*, Nanos/Pumilio complex maintains germline fate in migrating PGCs, halts cell cycle and inhibits apoptosis (Asaoka-Taguchi *et al*., 1999), while in *C. elegans*, *nanos1* and *nanos2* are required for maintenance and survival but not specification of PGCs, and Nanos3/Pumilio complex has a role in sperm-to-oocyte transition, but not in germ cell development (Ewen-Campen *et al*., 2010). *pum* (*pumilio*) genes were identified in zebrafish and rainbow trout but not yet in Nile tilapia. In Nile tilapia, three predicted *pum* genes were revealed by automated computational analysis by NCBI, but their molecular functions have not been studied yet.

Knockout of *nanos2* and *nanos3* genes resulted in complete loss of spermatogonia or germ cells both in mice (Tsuda *et al*., 2003) and Nile tilapia (Li *et al*., 2014), suggesting a conserved role in maintaining germ cells across phyla. However, there are four *nanos* genes in Nile tilapia, and their possible roles in PGCs are unknown. *nanos1* is duplicated in teleost fish genomes and named *nanos1a* and *nanos1b* in species including medaka and Nile tilapia (Aoki *et al*., 2009; Bellaiche *et al*., 2014). Consistent with the expression of *nanos1* in mice, both medaka *nanos1a* and *nanos1b* are expressed in the developing brain (Aoki *et al*., 2009). In medaka gonads, *nanos1a* is expressed in the gonadal somatic cells of both sexes, while *nanos1b* is not expressed in the gonads (Aoki *et al*., 2009). The tissue distribution of *nanos1a* and *nanos1b,* as well as ontogenic expression patterns, are lacking in Nile tilapia.

One of the XY gonadal marker genes, *nanos2*, is predominantly expressed in male germ cells and without them, mice cannot produce spermatogonia (Tsuda *et al*., 2003). In both zebrafish and medaka, *nanos2* was not detected in the early developmental stages of gonads; however, it was expressed in both oogonia and spermatogonia in adult gonads (Aoki *et al*., 2009; Beer and Draper, 2013). In rainbow trout, *in situ* hybridisation (ISH) of several germline-specific genes in maturing testis revealed stage-specific germ cell expression patterns. For example, *nanos2* is only expressed in some A spermatogonia, while *dazl* is expressed in all stages of germ cells (Bellaiche *et al*., 2014). In Nile tilapia testis, *nanos2* is expressed in type A spermatogonial cells, but not in type B cells (Fig. 1.6). *nanos2* can, therefore, be used as a type A spermatogonia marker in this species (Lacerda *et al*., 2013). In addition, knockout of *nanos2* in XY tilapia by CRISPR resulted in a loss of PGCs at 72 hpf and germ cell-deficient gonads at 90 dah (Li *et al*., 2014), implying that before 72 hpf, *nanos2* is transcribed zygotically and the *nanos2* transcripts are important in PGCs survival in Nile tilapia. However, it is still unknown when *nanos2* starts to be expressed during ontogenic development.
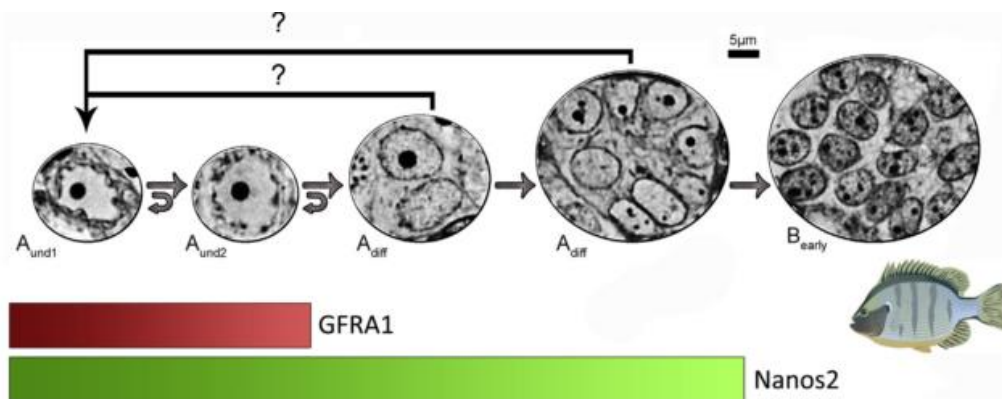
Figure 1.6. Schematic overview of protein expression regime of Gfra1 and Nanos2 in the different types of spermatogonial cells in Nile tilapia. Gfra1 expression (in red) is restricted to type A undifferentiated spermatogonia (A_{und1} and A_{und2}), while Nanos2 expression (in green) is observed in both A_{und} spermatogonia and type A differentiated spermatogonia (A_{diff}), but not in type B spermatogonia. Figure from Lacerda *et al*. (2013).

Both in zebrafish and medaka, maternally provided *nanos3* is located and expressed preferentially in PGCs, playing a role in maintenance and survival of PGCs, but not in the specification of PGCs in zebrafish (Beer and Draper, 2013; Herpin *et al*.,

2007; Köprunner *et al*., 2001). It is also shown that in zebrafish, Nanos3 protein interacts and regulates phosphorylation of myosin light chain II, which is important for the formation of pseudopodia leading to PGC migration during gastrulation (Blaser *et al*., 2006; Xu *et al*., 2010b). In Atlantic cod (*Gadus morhua*), expression of *nanos3* was detected from first cleavage division and ceased during early somitogenesis (Presslauer *et al*., 2012). In Nile tilapia, the expression of *nanos3* was reported in XX embryos at 36 - 72 hpf (quantitative PCR, qPCR) and was exclusively expressed in the ovary at 30 - 180 dah (ISH) (Li *et al*., 2014). Similar to *nanos2*, knockout of *nanos3* in XX tilapia by CRISPR resulted in the loss of PGCs at 72 hpf, germ cell-deficient gonads at 60 dah and female to male sex reversal (Li *et al*., 2014). Similar results were shown in zebrafish, where a mutation in *nanos3* resulted in the failure of maintaining oocyte production and loss of fertility (Draper *et al*., 2007). However, the molecular characteristics of *nanos3* in early development in Nile tilapia remain unknown.

- Small non-coding RNA (sncRNA)

sncRNA pathways participate in the post-transcriptional regulation of germ cells (de Mateo and Sassone-Corsi, 2014) including microRNAs (miRNAs, 19 - 25 bp) and piwi-interacting RNAs (piRNAs, mostly 24 - 34 bp) pathways (Saxe and Lin, 2011). miRNAs and Ago proteins which are components of microRNP are highly concentrated in germ granules in mice germ cells and they play an important role in spermatogenesis by post-transcriptional gene regulation (Kotaja *et al*., 2006). ATP-dependent DEAD-box RNA helicase, Vasa is exclusively expressed in the germ granules and involved in miRNA and piRNA pathways, essential for spermatogenesis (Kotaja *et al*., 2006). piRNAs are generated in a Dicer-independent manner and interact with Piwi proteins, playing an important role in silencing of retrotransposons in spermatogenesis (Aravin *et al*., 2007; Chuma and Nakano, 2013). The control of these mobile genetic elements is crucial to protect their genetic information in the germline genome (Kuramochi-Miyagawa *et al*., 2010). In both *Drosophila* and mice, animals lacking Piwi pathway resulted in sterile males implying an essential role of Piwi pathway in germline development (Klattenhoff and Theurkauf, 2008; Pillai and Chuma, 2012).

- *piwil-like 1* (*piwil1*) and *piwil2*

Piwi family members are components of gene regulatory network of germline cells, together with Tudor domain protein and DEAD box helicase, which maintain germ cell

fate and pluripotency of germline stem cells (Ewen-Campen *et al.*, 2010). The piwi family has two distinct domains: PAZ domain, an RNA binding motif, and Piwi domain, a similar structure of RNase H catalytic domain (Simon *et al.*, 2011). Piwi proteins are germline-specific argonautes, which combine with small RNA partners, piRNAs, and form RNA-induced silencing complexes (RISC) (Aravin *et al.*, 2007; Tolia and Joshua-Tor, 2007). The most well-established role of Piwi family proteins is silencing of transposable elements in the germline, showing that Piwi family mutations-induced germ cell loss is correlated with reduced piRNA levels and abnormal accumulation of transposable elements (Ewen-Campen *et al.*, 2010). In *Drosophila*, Piwi proteins appear to regulate the number and division of germline stem cells (GSCs) (Cox *et al.*, 2000). In mouse, *Miwi, Mili* and *Miwi2* (murine homologs of *piwil 1, 2* and *4*, respectively) are essential for spermatogenesis (Carmell *et al.*, 2007; Deng and Lin, 2002; Kuramochi-Miyagawa *et al.*, 2004).

Unlike mammals where piRNAs are testis-specific, in zebrafish, piRNAs are present in both testis and ovary, and many of them are derived from transposons, suggesting a role in transposon silencing (Houwing *et al.*, 2007). In zebrafish, Ziwi (zebrafish homolog of *piwil1*) protein was maternally provided and localised at cleavage planes in 4 cell stage as a GP component, and the loss of *ziwi* showed consecutive loss of germ cells due to apoptosis during early development, suggesting a role in the maintenance of germ cells (Houwing *et al.*, 2007). In addition, in zebrafish the *piwil2* homolog, *zili,* is known to be essential for germ cell differentiation and meiosis in the species (Houwing *et al.*, 2008). In medaka, *piwil1* and *piwil2* were also maternally deposited and expressed in PGCs in larvae, oogonia and mature oocytes in the ovary and spermatogonia and spermatocytes in adult testes (Zhao *et al.*, 2012). The important role of *piwil2* in gametogenesis of phenotypic males was also suggested in half-smooth tongue sole where *piwil2* transcripts are more abundant in testis than ovary and the gene is located on Z sex chromosome (ZZ/ZW-type) with lower CpG methylation levels than in female (Zhang *et al.*, 2014a). In Nile tilapia, there are also two *piwil* genes, *piwil1* and *piwil2*: both proteins are present in spermatocytes but not in mature sperm (Xiao *et al.*, 2013). Despite the apparent importance of Piwi family members in germline cells in diverse species, little is known about their molecular characteristics in fish.

Various genes including *vasa, dnd, nanos*, *pum* and *piwil* are known to be involved in translational repression in germ cells, which is thought to be key to the maintenance of the germline across animal phyla including zebrafish (Mickoleit *et al.*, 2011), *Xenopus*

(Lai and King, 2013), mouse (Carmell *et al*., 2007) and *Drosophila* (Asaoka-Taguchi *et al*., 1999; Cox *et al*., 2000). In Nile tilapia, to date, three *vasa* genes (Conte *et al*., 2017; Fujimura *et al*., 2011), four *nanos* (Bellaiche *et al*., 2014; Škugor *et al*., 2014a) and two *piwil* genes (Xiao *et al*., 2013) have been identified. Further *in silico* analysis of public databases by NCBI revealed predicted sequences for one *dnd1* as well as three *pum* genes. Overall, while evidence in vertebrates indicates pivotal roles of *vasa, dnd, nanos*, *pum* and *piwil* genes on PGCs and germline development, these genes have not been characterised in early ontogenic development of Nile tilapia.

1.3.3 Visualisation of PGCs

Prior to the identification of PGCs-specific molecular markers, alkaline phosphatase was used as an enzyme marker of PGCs (Molyneaux and Wylie, 2004). However, this method was not able to precisely label PGCs and was limited to specific stages of PGCs. More efficient ways to identify PGCs have been developed using PGCs-specific genes. To label PGCs, gene constructs containing reporter genes such as green fluorescence protein (GFP) or red fluorescence protein (RFP) conjugated with the promoters of germ cell markers such as *nanos3* or *vasa* were used in zebrafish, medaka and trout (Fan *et al*., 2008; Li *et al*., 2009; Yoshizaki *et al*., 2000). In addition, utilisation of 3'UTR of *nanos3* or *vasa* significantly improved the visualisation and detection of PGCs in teleosts (Lin *et al*., 2012; Nagasawa *et al*., 2013; Saito *et al*., 2011, 2006; Škugor *et al*., 2014a; Yoshizaki *et al*., 2005). In Nile tilapia, the microinjection of GFP-*vasa* 3'UTR mRNAs into zygotes was employed successfully to label PGCs in larvae (Li *et al*., 2014).

1.4 Apoptosis of PGCs

Apoptosis is a process of programmed cell death which is essential for normal development and maintenance of normal cellular homeostasis in metazoans (Danial and Korsmeyer, 2004). This programmed cell death plays also an important role in the elimination of mis-migrated PGCs. After specification of PGCs, some of PGCs fail to migrate into the genital ridge. The resultant ectopic PGCs undergo apoptosis, which is vital to prevent the formation of germ cell tumours as reported in mammals (Göbel *et al*., 2000; Schneider *et al*., 2001; Ueno *et al*., 2004). The apoptotic pathways are conserved in fish and mammals, although some differences can be found such as the lack of a C-

terminal region in the FADD (Fas-associated death domain) in teleosts (AnvariFar *et al*., 2017). While knowledge of these pathways is well advanced in mammals, very little is known in fish.

1.4.1 Intrinsic and extrinsic pathways of apoptosis

Once apoptosis has begun, the execution of the cell death is inevitable; therefore, the initiation of apoptosis is securely regulated (Böhm and Schild, 2003). Signalling pathways mediating apoptosis involve various molecules including Bcl-2 (B-cell lymphoma 2) family proteins, Caspases, cytochrome c, tumor antigen p53, Fas (death receptor) and FADD. There are two major apoptosis pathways; the intrinsic and extrinsic pathways, which are distinguished by the dependence on the accumulation of mitochondrial cytochrome c (Chauhan *et al*., 1997). The extrinsic pathway is an independent process from the accumulation of mitochondrial cytochrome c, and mediated by the binding between extracellular death ligands and cell-surface death receptor, which activates Caspase-8, resulting in executing the apoptosis of the cells (Wajant, 2002). On the other hand, Bcl-2 family proteins are closely involved in the regulation of the intrinsic apoptotic pathways. Bax (Bcl-2-associated X) protein up-regulates the permeability of the mitochondrial membrane, which enables the release of cytochrome c from mitochondria into the cytosol (Green and Reed, 1998; Kroemer *et al*., 2007). Then, the released cytochrome c binds to Apaf-1 (apoptotic peptidase activating factor-1) and Caspase-9 to activate Caspase-3, which results in cell death (Danial and Korsmeyer, 2004). In the intrinsic pathway, a change in the ratio between prosurvival and proapoptotic bcl-2 is the decisive step for the initiation of apoptosis (Jia *et al*., 2007; Oltval *et al*., 1993; Rucker *et al*., 2000).

1.4.2 Bcl-2 family genes

Bcl-2 family genes can be divided into two groups based on their roles, the proapoptotic role, *e.g. bax, bim* (*Bcl-2 homology 3-only protein*)*, box* (*Bcl-2-related ovarian killer*) and the prosurvival role, *e.g. bcl-xL* (*B-cell lymphoma-extra large*)*, bcl-2, mcl-1* (*myeloid cell leukemia 1*) (Cory and Adams, 2002). In mice, Rucker *et al*. (2000) revealed that a balance of Bcl-x (a cell survival factor) and Bax (a cell death factor) regulates PGC survival and apoptosis. In mice of both sexes, hypomorphic alleles of *bcl-x* induced severe germ cell loss and the concomitant removal of *bax* genes restored the survival of

germ cells (Rucker *et al*., 2000). Particularly, pro-apoptotic *bax* genes are known to be required for the intrinsic apoptosis of the ectopic germ cells (reviewed by Dicou and Perez-Polo, 2009). These results suggest that Bax or Bcl-x can be the potential targets for sterilisation of fish by inducing apoptosis of PGCs.

Teleosts species also possess Bcl-2 family genes comparable to mammalian homologs, but the studies are limited to a few species, including zebrafish (Kratz *et al*., 2006; Valentijn *et al*., 2008) or a few genes of Bcl-2 family such as *bax* (Ching *et al*., 2013). In zebrafish, the members of Bcl-2 family genes were identified and it showed that they are structural and functional homologs of most mammalian Bcl-2 family members by homology and phylogenetic analyses with human and mice and morpholino-mediated knockdown analysis (Kratz *et al*., 2006). Due to the genomic duplication in teleost, some of mammalian Bcl-2 family genes have two copies in zebrafish (Kratz *et al*., 2006). It has been shown that of the two zebrafish bax genes, zBax1 and zBax2, zBax2, but not zBax1, shares a conserved synteny with human Bax, suggesting a potential functional difference between the two paralogs (Kratz *et al*., 2006). However, whether their functions in germline lineage are the same in teleosts is unknown. Therefore, further studies are required to determine their functions and understand the mechanisms that regulates the survival and apoptosis of PGCs.

1.4.3 Effect of high temperature on the survival of germ cells

In poikilotherms, heat stress is a major environmental stressor and it can cause germ cell loss through the stimulation of apoptotic pathways in fish (AnvariFar *et al*., 2017). Therefore, high temperature treatment can be exploited to induce apoptosis of germ cells to understand the underlying mechanism of germ cell death/survival in fish.

In mice, rat and monkey, exposure to elevated temperature caused stage-specific apoptosis of germ cells in testes (Lue *et al*., 2002, 1999; Yin *et al*., 1997). In mice, *p53* is found in high concentrations in the testis and high temperature alters the expression of *p53*, resulting in germ cell loss (Socher *et al*., 1997). In addition, in teleosts it has been shown that water temperature above the natural range for the species can hinder gonadal development, causing germ cell loss (de Alvarenga and de França, 2009; Ito *et al*., 2003; Soria *et al*., 2008; Strüssmann *et al*., 1998; Uchida *et al*., 2004). However, it has not been revealed in these cases whether apoptotic genes are involved in temperature-mediated germ cell loss.

Baroiller *et al*. (2009) suggested apoptosis and/or proliferation of PGCs caused by temperature during the labile period of sex differentiation could be a critical factor for sex reversal in fish. Under high temperature, early diplotene stage oocytes died in all genetically female zebrafish exposed to 37°C between 15 - 25 dah, showing 100% masculinisation (Uchida *et al*., 2004). Likewise, exposure to 32°C during early gonadal development induced germ cells degeneration and masculinisation of ovarian somatic cells, with a concomitant increase of *dmrt1* expression in pufferfish (*Takifugu rubripes*) (Lee *et al*., 2009). In medaka, high temperature (33°C) inhibited proliferation of germ cells and induced female-to-male sex reversal associated with an increase in plasma cortisol (Hayashi *et al*., 2010). Exposure to the high temperature of 29°C induced degeneration of germ cells showing clear pyknosis or eosinophilia in both larval and adult stages of male Patagonian pejerrey (Ito *et al*., 2003; Soria *et al*., 2008).

Moreover, complete sterility was obtained by exposure to the upper incipient lethal temperature in female *P. hatcheri* and *O. bonariensis* at 27 - 28.5 and 29°C, respectively (Strüssmann *et al*., 1998). Recently, it has been reported that a high temperature treatment of 37°C for 40 - 50 days from 3 dah induced apparent sterility in female Nile tilapia (Chitralada strain) and male Mozambique tilapia (Nakamura *et al*., 2015; Pandit *et al*., 2015). In addition, exposure of sexually mature adult male Nile tilapia to 30 - 35°C for a week showed induced germ cell differentiation with lower Sertoli and Leydig cell proliferation compared to fish exposed to 20 - 25°C, but the number of germ cells was not affected (de Alvarenga and de França, 2009; Lacerda *et al*., 2006). These data suggest that there are species-specific thermal sensitivities in germ cell survival and the incipient lethal temperature can effectively induce germ cell death. While evidence clearly showed the temperature-specific effects of heat treatments on gonadal germ cells and the overall suppression of gametogenesis, the underlying mechanisms leading to germ cell death remains unclear. In subadult pejerrey, the exposure to high temperature induces apoptosis in both somatic and germ cells with significant increases of Caspase-3 activity in both genders (Ito *et al*., 2008). The severity of heat-induced apoptosis was proportional to the magnitude of the thermal stress and males tended to be more sensitive to thermal stress than females (Ito *et al*., 2008). Still, the underlying mechanism of heat-induced apoptosis of gonads is limited in teleosts. Oltval *et al*. (1993) suggested that the balance between proapoptotic and prosurvival Bcl-2 family proteins can determine survival or death of cells, as demonstrated in mouse PGCs where the ratio of Bcl-x and Bax regulates the

survival of PGCs and apoptosis (Rucker *et al*., 2000). In teleosts, however, it is unknown which Bcl-2 family members are involved in temperature-mediated germ apoptosis.

1.5 CRISPR/Cas9 genome editing technique for gene functional analysis in teleosts

1.5.1 CRISPR/Cas system

Over the past decade, genome editing techniques have rapidly improved in precision and efficiency. The first generation techniques included zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) in which DNA binding domains are programmable for site-specific DNA cleavage (Gaj *et al*., 2013). ZFNs are composed of the nonspecific DNA cleavage domain from the FokI endonuclease and sequence-specific DNA binding zinc-finger domains (ZFDs) recognising a nucleotide triplet of the target DNA (Doyon *et al*., 2008; Kim *et al*., 1996; Wolfe *et al*., 2000). TALENs also use the nonspecific DNA cleavage domain from the FokI endonuclease, but the sequence-specific DNA binding domains are different from ZFNs: they are called transcription activator-like effectors (TALEs), and each TALE recognises a single base pair rather than a triplet in the case of ZFD (Bogdanove and Voytas, 2011). Both ZFNs and TALENs induce double-strand breaks (DSBs) on the targets, which are repaired by DNA damage response pathways, enabling gene functional analysis. Even though both ZFNs and TALENs use the same nuclease from FokI, TALENs typically reports a higher mutation rate, lower off-target effects, lower cytotoxicity and more targetable sites (ZFNs, one per approximately 100 bp; TALENs, at least one per base pair), but the size is approximately three times bigger than ZFNs (reviewed by Kim and Kim, 2014). More recently, the second generation of genome editing technique called CRISPR/Cas came along, and all three genome editing techniques opened a new era for functional gene study.

CRISPR/Cas is an adaptive immune system found in approximately 50% of bacteria and 90% of archaea to protect themselves from viral infection (Lander, 2016; Makarova *et al*., 2015). Over the past few years, researchers found that this system can be engineered to edit any nucleic acid sequences as well as viral nucleic acids sequences by designing the guide RNA to match the target, which can be done not just *in vitro*, but also within the nucleus of a living cell (Lander, 2016). The CRISPR-Cas system falls into two classes: Class 1 systems use a complex of multiple protein effectors to degrade foreign nucleic acids, while Class 2 systems use a single-component large protein

effector (Makarova *et al*., 2015) (Table 1.1). In the Class 1, type I is the most common and diversified CRISPR/Cas system and type III is more frequently found in archaea than bacteria (Koonin *et al*., 2017) (Table 1.1). The Class 2 possesses a single multidomain protein effector.

Table 1.1. Classification of CRISPR systems (Koonin *et al*., 2017; Lander, 2016)

Class	Type	Hallmarks/function	Examples effector	Target	Example organism	References
1	I	Multisubunit effector complex; containing Cas3 which is single-stranded DNA-stimulated superfamily 2 helicase	Cascade*	DNA	*Escherichia coli*	(Brouns *et al*., 2008; Sinkunas *et al*., 2011)
	III	Multisubunit effector complex; containing Cas10 which possesses two cyclase-like Palm domains	Cas10	DNA/RNA	*Staphylococcus epidermidis, Pyrococcus furiosus*	(Hale *et al*., 2009; Marraffini and Sontheimer, 2008; Zhu and Ye, 2012)
2	II	Single protein effector; tracrRNA; Cas9 ensures the acquisition of functional spacers during CRISRP-Cas adaptation	Cas9	DNA	*S. thermophilus, S. pyogenes*	(Gasiunas *et al*., 2012; Heler *et al*., 2015; Jinek *et al*., 2012)
	V	Single protein effector; Cas12a is guided by single-RNA (lacking tracrRNA) while Cas12b requires both crRNA and tracrRNA	Cas12a (Cpf1), Cas12b (C2c1)	DNA	*Francisella novicida, Alicyclobacillus acidoterrestris*	(Liu *et al*., 2017; Zetsche *et al*., 2015)
	VI	RNA-guided RNase; once activated by the cognate target RNA, it cleaves collateral RNA in a non-specific manner; posing cellular toxicity	Cas13a (C2c2), Cas13b	RNA	*Leptotrichia shahii, Prevotella buccae*	(Abudayyeh *et al*., 2016; Cox *et al*., 2017; Smargon *et al*., 2017)

*Cascade, CRISPR-associated complex for antiviral defence.

Cas9 belongs to Class 2/type II and consists of a single protein effector derived from either *Streptococcus pyogenes* or *S. thermophilus* and it has two nuclease motifs, RuvC and HNH, producing DSBs at 3 nucleotides upstream of the protospacer adjacent motif (PAM) (Deveau *et al*., 2008; Makarova *et al*., 2006). When the bacterium detects the presence of viral DNA, it produces two types of short RNAs, one of which is transcribed from the CRISPR locus with the sequence matching that of the invading virus, called CRISPR RNA (crRNA), and the other is complementary to the repeat sequences in pre-crRNA, called transactivating crRNA (tracrRNA), forming an active complex by binding with crRNA (Deltcheva *et al*., 2011; Jinek *et al*., 2012). Inside the nucleus, the resulting complex will recognise and bind to PAM, and then, helicase the dsDNA and hybridise with guide RNA. When the match is complete, the target DNA is cut by molecular scissors, RuvC and HNH, in Cas9 effector. To edit genome using CRISPR/Cas9 system, crRNA, tracrRNA and Cas9 were required originally, but now only two components are needed by merging crRNA and tracrRNA into a single guide RNA (sgRNA) (Jinek *et al*., 2012).

The Class 2 CRISPR-Cas system includes Cas12a (Cpf1) in type V as well as the widely used endonuclease, Cas9 in type II. Unlike Cas9, Cas12a does not require tracrRNA for target cleavage (Zetsche *et al*., 2015). The common feature of types II and V is the presence of a RuvC-like endonuclease domain that adopts the RNase H fold (Koonin *et al*., 2017). Class 2 type VI is an RNA-targeting CRISPR/Cas system and the type VI effector, Cas13 enzyme, has two HEPN (higher eukaryotes and prokaryotes nucleotide-binding) RNase domains that mediate precise RNA cleavage with a preference for targets with protospacer flanking site motifs (Abudayyeh *et al*., 2016; Shmakov *et al*., 2015). Interestingly, once activated by the target RNA, Cas13 cleaves RNA in a non-sequence-specific manner (Abudayyeh *et al*., 2016; Smargon *et al*., 2017). In addition, the HEPN RNase in Cas13 protein is the toxin domain, which induces cellular toxicity, potentially resulting in dormancy or programmed cell death in bacteria (Abudayyeh *et al*., 2016). Due to the simplicity in usage and precise and efficient gene editing, engineered Cas9 endonuclease is the most widely used protein effector as a genome editing tool among diverse protein effectors (Table 1.1). Yet, much more novel effector proteins await to be elucidated and it will enable various modes of editing target genes, *e.g.* dsDNA, ssDNA, dsRNA or ssRNA; sequence-specific or non-specific cleavages of genes.

Among various Cas effectors, so far, Cas9 has been successfully adopted in genome editing studies in fish. Since the CRISPR/Cas9 system has proven to be a simple, efficient and reliable method for targeted mutagenesis in zebrafish (Jao *et al*., 2013), it has been used in various species including medaka (Ansai and Kinoshita, 2014), Nile tilapia (Li *et al*., 2014), Southern catfish (*Silurus meridionalis*) (Li *et al*., 2016), channel catfish (Khalil *et al*., 2017), common carp (Zhong *et al*., 2016), Atlantic salmon (Edvardsen *et al*., 2014), killifish (*Fundulus heteroclitus*) (Aluru *et al*., 2015) and rohu (*Labeo rohita*) (Chakrapani *et al*., 2016). The easy programmability and high mutation efficiency of this system boosted gene functional analyses to understand biological functions such as sex determination pathway in Nile tilapia (Li and Wang, 2017). Nile tilapia has clear advantages as a "model organism" for genome editing technology (Maclean *et al*., 2002). First of all, it has a short generation time of 6 - 7 months to reach maturity and their eggs are constantly available. Tilapia embryos are semi-transparent, so some reporter genes can be detected to some extent without sectioning or sacrificing (Maclean *et al*., 2002). More importantly, the genome of Nile tilapia is fully sequenced and accessible (Accession: PRJNA59571) (Brawand *et al*., 2014). In contrast, there is an important downside that must be acknowledged, the eggs of the species are less easily injected than those of zebrafish, medaka and carp due to harder chorion. However, the eggs can be injected either via the micropyle or through the chorion but then each egg must be individually positioned for optimal needle entry which can be time consuming.

One of the significant advantages of CRISPR/Cas9 system is the possibility to target multiple genes at once, which is especially beneficial for studying complex diseases or traits/systems controlled by multiple genes acting together (Cong *et al*., 2013). In addition, the high efficiency of CRISPR/Cas9 mediated genome editing can produce biallelic mutants in founder (F0) individuals, allowing a direct phenotypic analysis without generating F1 and F2 generations (Jao *et al*., 2013). This enables the investigation of gene function in species which have a long generation time such as salmonids or genes associated with reproductive competence, *i.e.* sterility where successful disruption would mean that subsequent progeny production would not be possible. However, the lack of reliable and efficient screening methods for F0 remains a problem.

1.5.2 Issues in mutation screening methods

Once the target gene has been cut by the effector complex, cells try to repair the cut. Most of the cut sites are repaired by an error-prone non-homologous end joining (NHEJ) method, which causes insertion or deletion (indel) at the cut sites, leading to mutations that can disable the gene function, allowing researchers to understand its function (Rodgers and McVey, 2016). On the other hand, the homology-directed repair (HDR) method can edit the cut site precisely by adding a donor DNA which pairs up with the cut ends, recombining and replacing the original sequence with the new donor sequence (Huang *et al*., 2011). HDR leads to precise genome editing, including targeted gene insertion, correction and point mutagenesis (Kim and Kim, 2014). However, NHEJ repair pathway is more dominant than HDR following DSBs-induced by CRISPR/Cas9 (Mali *et al*., 2013; Wang *et al*., 2013; Yang *et al*., 2013), as HDR occurs only in synthesis (S) and second gap (G2) phase while NHEJ occurs in entire cell cycle (Deriano and Roth, 2013). Recent studies attempted to improve HDR efficiency in various ways such as controlling the cell cycle (Heyer *et al*., 2010; Lin *et al*., 2014; Orthwein *et al*., 2014), inhibiting the NHEJ process (Li *et al*., 2017b; Maruyama *et al*., 2015) or modifying the donor DNA by phosphorothioate (Liang *et al*., 2017). In addition, usages of HDR enhancer molecules (Shao *et al*., 2017; Song *et al*., 2016), nucleofection of Cas9 RNP complex (Kim *et al*., 2014b; Lin *et al*., 2014) or ssDNA (Richardson *et al*., 2016) were also tested to enhance the efficiency of HDR. However, the efficiency of HDR still needs to be improved for applications such as functional gene analysis in injected animals (F0).

These endogenous repair methods of programmable nuclease-induced strand breaks can be performed in cultured cells or in a fertilised egg, producing mutated cells or animals with targeted mutations. Unlike mutated cells in which the cell genotype is relatively simple such as homozygous or heterozygous mono- or biallelic mutations, the founder (F0) mutant animals created by CRISPR/Cas9 and NHEJ DNA repair system possess high mosaicism (Edvardsen *et al*., 2014; Oliver *et al*., 2015; Yen *et al*., 2014) as the gene editing is unlikely to be completed at the 1-cell stage (Samarut *et al*., 2016). CRISPR/Cas9-induced mutations are different from those produced by conventional gene-targeting approaches which incorporate exogenous DNA and/or allele-specific sequence that can be exploited to facilitate genotyping, *e.g*. using primers designed on the exogenous DNA sequence inserted (KC *et al*., 2016). One of the bottlenecks in mutagenesis studies using the CRISPR/Cas9 system is the lack of comprehensive and

efficient screening methods for the F0 mutants. The small indels or substitution created by NHEJ DNA repair are often not easily resolved with traditional genotyping strategies, requiring other strategies which are sensitive enough to detect small indels or substitutions (Kim and Kim, 2014). On the other hand, the large indels can be overlooked in some screening methods based on PCR, in which the deletion is beyond the primer sites. To overcome this bottleneck, high-throughput time- and cost-effective genotyping methods with a wide detection coverage of various mutations induced by CRISPR/Cas9 and NHEJ should be validated.

1.5.2.1 Comparison of mutant screening methods

1.5.2.1.1 Mismatch detection assays

The typical three steps of mismatch detection assays are: (1) amplification of the target site by PCR, (2) formation of heteroduplex DNAs by denaturing and reannealing the DNA and (3) detection of the heteroduplex, which is a relatively simple, rapid and cost-effective assay (Zischewski *et al.*, 2017). Although mismatch detection is often used as a preliminary screening approach to identify mutants, it hardly reveals details of the mutation structure, therefore, further analysis using more informative sequencing-based methods should be followed (Zischewski *et al.*, 2017). Furthermore, if the targeted locus is highly polymorphic, it is difficult to interpret the results due to the presence of heteroduplex DNAs formed with different WT alleles (Kim *et al.*, 2014a). Another potential problem with PCR based mutation screening methods is that large deletions that extend beyond the boundaries of the PCR amplicon are not detected, and large insertions amplified less efficiently than small mutations may, therefore, not be identified (Zischewski *et al.*, 2017). Mismatch detection assays include mismatch cleavage assays such as T7 Endonuclease 1 (T7E1) and SURVEYOR nuclease, melt curve analysis, high resolution melt analysis (HRMA) and heteroduplex mobility assay.

The potential problems and limitations of mismatch cleavage assay such as T7E1 and SURVEYOR were reviewed by Zischewski *et al.* (2017). According to these authors, these assays can underestimate the mutation frequency due to the preferential cleavage properties of each enzyme and homozygous mutations cannot be detected unless adding WT DNA to the PCR step. In addition, the enzyme reactions can give false negative results caused by the incomplete digestion of mismatched DNA fragments.

Melt curve analysis including HRMA are fast, high-throughput and sensitive methods (96- or 384-well microtiter plates). Unlike typical qPCR melt curve analysis, HRMA requires a saturating fluorescent dye which is suitable to collect fluorescence data acquisition at $0.04 - 0.05°C$ increments in melt curve analysis. The result of both analyses can be shown as melting temperature shifts or the shape of the melt curves: homoduplexes from homozygous mutation cause a temperature shift, while heteroduplexes from heterozygous mutations modify the shape of the melt curve (Chateigner-Boutin and Small, 2007). The shorter the amplicon, the greater the resolution and the less detection of polymorphism and also larger indels (Dahlem *et al*., 2012). The setup cost of HRMA can be high unless the existing qPCR machine is compatible with HRMA setting and the software, but once the equipment is ready, the cost per analysis is low (Dahlem *et al*., 2012).

Heteroduplex mobility assay detects mutations by using the differential migration between heteroduplex and homoduplex DNA in polyacrylamide gel electrophoresis (PAGE), as heteroduplex DNA migrates at a significantly slower rate than homoduplex DNA (Zhu *et al*., 2014; Zischewski *et al*., 2017). However, similar to HRMA, the PAGE assay also resolves only small amplicons, so large indels cannot be detected, and the sensitivity of the PAGE assay across the spectrum of indels remains to be verified (Shui *et al*., 2016; Zhu *et al*., 2014).

1.5.2.1.2 Restriction fragment length polymorphism (RFLP)

The RFLP assay involves the amplification of the target site, restriction enzyme digestion, followed by gel electrophoresis to measure the intensities of the cleaved and uncleaved DNA fragments, which are used to estimate indel frequency (Ran *et al*., 2013). Unlike the mismatch detection assay, RFLP does not involve heteroduplex formation, but detects homozygous mutants and is not affected by sequence polymorphisms near the target site. However, the prerequisite of this assay is that the target site of CRISPR/Cas9 is designed on a restriction enzyme site, which substantially limits the availability of potential target sites. In addition, some mutations, which do not disrupt the restriction site, will not be detected, and the enzyme reaction-based assay can give false positive results caused by the incomplete digestion.

1.5.2.1.3 CRISPR/Cas-derived RNA-guided engineered nucleases (RGENs)

Kim *et al*. (2014a) suggested the use of the identical sgRNA used for mutagenesis to detect the unmodified allele by cutting with Cas9. It has the advantage over RFLP and T7E1 as it does not require specific restriction enzyme sites on the target region and it is free from polymorphism-caused false positive results. In addition, unlike the mismatch cleavage assays, RGENs can distinguish biallelic homozygous mutants from WT and monoallelic heterozygous from biallelic heterozygous mutants (Kim *et al*., 2014a). However, this assay also can give false positive results caused by the incomplete enzyme digestion.

1.5.2.1.4. Loss of a primer binding site

This method is a PCR-based protocol for detecting indel mutations. It requires 2 pairs of primers for each target locus, with one putative amplicon extending beyond the putative indel site and the other overlapping it (Yu *et al*., 2014). The PCR products are resolved by electrophoresis and used for estimation of indel frequency by comparing the intensities of the two PCR bands. It is relatively fast and cheap, but the major problem is that the primer can tolerate mismatches on the target site, which can cause underestimation of indel frequency. This method has a low sensitivity, so it is not suitable if the efficiency of CRISPR/Cas9 is expected to be low (Zischewski *et al*., 2017).

1.5.2.1.5. Sequencing (Sanger sequencing and next generation sequencing)

Sequencing methods provide sequence details which are possibly the most informative aproach for the identification of mutations. Sanger sequencing is one of the most widely used genotyping methods, which involves cloning and sequencing of 50-100 individual clones, providing a snapshot of the characterisation of the indel spectrum and a confirmation of some mutant sequences, but it is a highly labour-intensive and low-throughput method (Brinkman *et al*., 2014). Instead of sequencing each clone, it is also possible to sequence the pooled PCR amplicon using Sanger sequencing and then analyse the sequence reads using bioinformatic tools such as TIDE (tracking of indels by decomposition) (http://tide.nki.nl) (Brinkman *et al*., 2014). It reduces the workload of the conventional sequencing method, but it is noted that the result of TIDE is only reliable when the sequence read quality is good (Zischewski *et al*., 2017). In contrast, next generation sequencing (NGS) of library pools of PCR products provides comprehensive sequence detail and a characterisation of the mutation spectrum in a high-throughput

fashion, and sequencing costs can be reduced if a large number of samples are multiplexed (Brinkman *et al*., 2014). The weakness of NGS is that the reads are relatively short so that larger indels spanning over the amplicon size cannot be detected. Hendel *et al*. (2014) developed single molecule real-time (SMRT) DNA sequencing, which is able to read 8.5 kb in average, but the cost remains high and the sensitivity is lower than that of other sequencing platforms (Zischewski *et al*., 2017).

1.5.2.1.6. Fluorescent PCR capillary gel electrophoresis (Fragment analysis)

This method employs fluorophore-labelled primers for PCR reaction and the indel mutations are revealed by their different mobilities compared to the WT amplicon in capillary gel electrophoresis (Ramlee *et al*., 2015; Yang *et al*., 2015). This method is relatively fast, sensitive and high-throughput (96-well microtiter plates) like HRMA and NGS. Unlike HRMA, fragment analysis is able to reveal the distribution and size of indels in the mutant sample and is cheaper than NGS. The downsides of this method are: (1) the detail of the sequence remains unknown, (2) mutations which do not change the length are undetectable and (3) overestimation of large indels (>30 bp) has been reported (Ramlee *et al*., 2015). Both Ramlee *et al*. (2015) and Yang *et al*. (2015) reported that this assay can reliably detect 1 bp difference, thus it can identify frame-shift mutations to some extent, but the setup costs can be expensive for this analysis.

1.5.2.1.7 Summary

None of the techniques listed above is perfect, with each of them presenting some limitations hampering a reliable and comprehensive characterisation of the mutation (Table 1.2). To date, the validation of current mutant screening methods was mainly investigated in mutant cell-line, which has a different spectrum of mutation compared to the genotype of F0 mutant animals. However, the founder animals produced by genome editing technique such as ZFN, TALEN or CRISPR/Cas9 are highly mosaic in comparison to mutated cell-lines which have relatively simple modes of inheritance (either WT, homozygous, heterozygous or compound heterozygous mutations) with a maximum of two differently edited sequences in their target site (one per each allele). There is a clear knowledge gap in the selection of reliable mutant screening methods for F0 animals. Therefore, the ability and limitation of each method should be validated in F0 through a comprehensive comparison of various mutant screening methods to reveal

the effective and efficient high-throughput screening method for CRISPR/ Cas9-mediated F0 mutants harbouring indels in targeted gene.

Table 1.2. Overview of methods for the detection of on-target mutations induced by sequence-specific nucleases. Adapted and modified from Zischewski *et al*. (2017)

Methods	Coverage spanning on-target site	Reported sensitivity	Determination of mutation type?	Cost[a]	Throughput	Limitations	References
Mismatch cleavage assay	300 – 1000 bp	Moderate	No	$	Moderate	Not accurate for high mutation rate samples	(Kim *et al*., 2014a; Ran *et al*., 2013; Vouillot *et al*., 2015; Zhu *et al*., 2014)
HRMA	50 – 300 bp	Moderate	No	$	High	Misses larger indels than amplicon	(Dahlem *et al*., 2012; Wang *et al*., 2015)
Melt curve analysis	50 – 300 bp	Moderate	No	$	High	Misses larger indels than amplicon and small indels	(D'Agostino *et al*., 2016)
Heteroduplex mobility assay by PAGE	50 – 500 bp	High to moderate	No	$	Moderate	Misses larger indels than amplicon	(Zhu *et al*., 2014)
RFLP	300 – 1000 bp	Moderate to low	No	$	Moderate	Availability of restriction site on target site	(Jin *et al*., 2012; Ran *et al*., 2013)
RGENs	300 – 1000 bp	Moderate	No	$	Moderate	Misses single substitution	(Kim *et al*., 2014a)
Loss of primer binding site by qPCR	50 – 300 bp	Low	No	$	High	Misses substitutions and larger indels than amplicon	(Yu *et al*., 2014)
Sanger sequencing	<1000 bp	High	Yes	$$	Low	Costly, labour-intensive	(Brinkman *et al*., 2014)
NGS	<550 bp	High	Yes	$$$$	High	Misses larger indels than amplicon	(Hendel *et al*., 2015)
Fluorescent PCR-capillary gel electrophoresis	<500 bp	High	Length (bp) change	$$	High	Misses substitutions and larger indels than amplicon	(Ramlee *et al*., 2015; Yang *et al*., 2015)

[a] Estimated cost per assay. $, <1 US$; $$, <5 US$; $$$, >100 US$; $$$$, >500

1.6 Experimental Aims

The Nile tilapia is one of the major aquaculture species across the globe. It is also an important laboratory species in which the whole genome has been sequenced and is accessible. In addition, Nile tilapia has a relatively short generation time and zygotes can be available all year round (Guyon *et al*., 2012; Soler *et al*., 2010). One of the main challenges related to the farming of this species is the uncontrolled reproduction in mixed sex populations which results in the production of inbred progenies, too high densities and small sized fish at harvest due to their size (Phelps and Popma, 2000). As male tilapia grow faster than female, all male monosex tilapia is desired in tilapia farming (Dixon, 1994). It is currently achieved by either direct hormonal treatment to induce masculinisation (Little and Hulata, 2000), which is banned in several major producing countries (although not enforced in most cases) (Conte *et al*., 2017) or via indirect hormonal treatment through the use of YY supermales (Beardmore *et al*., 2001), which is commercially available but a difficulty in keeping the broodstock is a hindrance to the wide usage. In addition, both methods are not free from the negative impact on the environment posed by escapees. A new means of controlling reproduction would, therefore, be desired. To do so, it is important to understand the genetic basis of gonadal development in current aquaculture stocks to develop alternative sterilisation methods.

The overall research aim was to investigate the genes involved in the survival of germ cells and subsequently conduct a functional analysis of candidate genes to ultimately provide the basis for the development of a novel sterilisation technique. Therefore, the first experimental phase of the thesis aimed to find suitable target genes which play an important role in the survival or maintenance of germ cells. In **Chapter 2**, ontogenic expression and tissue distribution of 11 candidate genes (*nanos1a, nanos1b, nanos2, nanos3, piwil1, piwil2, dnd, vasa, pum1, pum2* and *pum3*) were investigated to screen maternally provided germline-specific genes. Furthermore, in **Chapter 3**, expression of Bcl-2 family genes (*bcl-xLa, bcl-xLb, baxa, baxl, bcl-2* and *mcl-1a*) as well as germline-specific genes (*nanos2, nanos3, piwil1, piwil2* and *gfra1*) were explored under heat stress (36 - 37°C) to elucidate major regulator genes in apoptosis and survival in testicular germ cells and potentially expand the suitable target genes for sterility. In addition, analyses of plasma sex steroid levels, testis histology and sperm quality including sperm density, motility and fertilisation rate were carried out to study the

impact of high temperature (36 - 37°C) on reproduction in juvenile Nile tilapia. Based on the results from Chapters 2 and 3, the candidate gene *piwil2* was chosen for further functional analysis using the genome-editing technique CRISPR/Cas9. However, given the number of variables that require optimisation with such a technique an iterative approach was required to address specific methodological challenges in series. **Chapter 4** was, therefore, dedicated to the optimisation of both the microinjection setup for newly fertilised tilapia eggs as well as the CRISPR/Cas9 system targeting *piwil2* with the aim being to generate a reliable and effective mutation induction methodology. Subsequently, in **Chapter 5**, *piwil2* KO tilapia were produced to investigate the impact on the survival of PGCs with an emphasis on characterising the physiological impacts on PGCs during early development. Finally, through large-scale testing of mutant screening methods (T7E1, RGENs, HRMA, fragment analysis and NGS), the relationship between the observed phenotypes and apparent complex genotypes of CRISPR/Cas9 induced *piwil2* KO tilapia was explored in **Chapter 6**.

CHAPTER 2

EXPRESSION PATTERN OF *NANOS*, *PIWIL, DND, VASA* AND *PUM* GENES DURING ONTOGENIC DEVELOPMENT IN NILE TILAPIA *OREOCHROMIS NILOTICUS*

YH. Jin, A. Davie, H. Migaud

Institute of Aquaculture, Faculty of Natural Sciences, University of Stirling, UK

2.1 Introduction

PGCs are set aside from somatic cells, migrate into prospective gonadal sites during early development and give rise to germ cell lineage (Cinalli *et al*., 2008). There is a growing interest within the aquaculture sector to explore new means to induce sterility through disruption of the normal PGC development (Wargelius *et al*., 2016; Wong and Zohar, 2015a; Zhang *et al*., 2015a). If successful, PGC disrupted animals would not sexually mature which has a range of economic as well as environmental benefits. However, before such a technology can be developed and applied in a given species, it is important to understand how PGCs are specified and maintain their germ cell fate.

PGCs are specified by either maternally derived GP as shown in fish, frogs, fruit flies and nematode worms or inductive signalling from surrounding tissues such as BMP signalling during early embryogenesis as shown in most mammals, axolotls and turtles (Extavour and Akam, 2003; Hogan, 1996; Raz, 2003). The maternally deposited GP consists of maternal mRNAs and proteins of germ cell-specific genes, and they are found in PGCs and shown to have essential roles in PGC specification, germ cell development and/or maintenance (Houwing *et al*., 2007; Knaut *et al*., 2000; Kosaka *et al*., 2007), but their functions are unclear in many teleost species. The application of PGC ablation methods in Nile tilapia (*O. niloticus*) requires first the identification of germ cell-specific genes that are maternally deposited and prioritising the target genes.

Many GP components are RNA-binding proteins or RNAs, which are required to maintain germ cell fate and totipotency, otherwise, they become somatic cells (Cinalli *et al*., 2008; Lai and King, 2013; Tada *et al*., 2012). *Nanos*, *pum*, *vasa, dnd* and *piwil* genes are known to be involved in translational repression in germ cells, which is thought to be key to the maintenance of the germline across animal phyla including zebrafish (Mickoleit *et al*., 2011), *Xenopus* (Lai and King, 2013), mouse (Carmell *et al*., 2007) and *Drosophila* (Asaoka-Taguchi *et al*., 1999; Cox *et al*., 2000).

In Nile tilapia, to date four *nanos* genes (Bellaiche *et al*., 2014; Škugor *et al*., 2014a), two *piwil* genes (Xiao *et al*., 2013) and three *vasa* genes (Conte *et al*., 2017; Fujimura *et al*., 2011) have been identified. Further *in silico* analysis of public databases by NCBI revealed predicted sequences for one *dnd1* as well as three *pum* genes. The predicted Nile tilapia Nanos proteins contain two CCHC zinc-finger domains which bind to RNA with low sequence specificity (Bellaiche *et al*., 2014). In Nile tilapia, Kobayashi (2010) reported that *nanos1* showed a sexually dimorphic expression in the sex

differentiation period, being expressed in oogenic meiotic cells. Following the publication of this work, however, the target studied (accession No. AB453384.1) has retrospectively been reclassified as *nanos3* in accordance with Bellaiche *et al.* (2014) and Škugor *et al.* (2014a). Knock-out (KO) of *nanos2* and *nanos3* genes resulted in complete loss of spermatogonia or germ cells both in mice (Tsuda *et al.*, 2003) and Nile tilapia (Li *et al.*, 2014), suggesting a conserved role in maintaining germ cells across phyla. However, there are four *nanos* genes in Nile tilapia, and their possible roles in PGCs are unknown.

Piwi family members are components of gene regulatory network of germline cells, together with Tudor domain protein and DEAD box helicase, which maintain germ cell fate and pluripotency of germline stem cells (Ewen-Campen *et al.*, 2010). Piwi proteins are germline-specific Argonautes, which combine with small RNA partners termed piRNAs and form RISC (Aravin *et al.*, 2007; Tolia and Joshua-Tor, 2007). In *Drosophila*, Piwi proteins appear to regulate the number and division of germline stem cells (GSCs) (Cox *et al.*, 2000). In mouse, *Miwi, Mili* and *Miwi2* are essential for spermatogenesis (Carmell *et al.*, 2007; Deng and Lin, 2002; Kuramochi-Miyagawa *et al.*, 2004). Unlike mammals where piRNAs are testis-specific, in zebrafish, the piRNAs exist in both testis and ovary, and many of them are derived from transposons, implicating a role in transposon silencing (Houwing *et al.*, 2007). In zebrafish, the loss of *ziwi* showed consecutive loss of germ cells due to apoptosis during early development, suggesting a role in the maintenance of germ cells (Houwing *et al.*, 2007). In addition, *zili*, is known to be essential for germ cell differentiation and meiosis in this species (Houwing *et al.*, 2008). In Nile tilapia, there are also two *piwil* genes (*piwil1* and *piwil2*); both proteins are present in spermatocytes but not in mature sperm (Xiao *et al.*, 2013). Despite the apparent importance of Piwi family members in germline cells in diverse species, little is known about their molecular characteristics in fish.

Dnd is an RNA-binding protein, which protects germ cell-specific RNAs from microRNA mediated degradation by binding to the 3'UTRs of germ cell-specific genes in human cell and zebrafish PGCs (Kedde *et al.*, 2007). Vasa belongs to an ATP-dependent RNA helicase of the DEAD box (Asp-Glu-Ala-Asp) family, and it is one of the well-known components of GP (Hay *et al.*, 1988; Yoon *et al.*, 1997).

The Puf domain of the RNA binding protein Pumilio provides the specificity of Nanos by binding specific target sequences through interacting with Nanos proteins, and the Nanos/Pumilio complex represses translation of target RNAs by deadenylation

(Ewen-Campen *et al*., 2010). In *Drosophila*, maternally derived Nos/Pum complex which binds to Pumilio binding element in the 3'UTRs of target mRNAs is essential to maintain germline cells (Asaoka-Taguchi *et al*., 1999). In zebrafish, PGCs-specific knockdown of a novel *puf-A* gene, clustered separately from *pum1* and *pum2*, caused the reduction of PGCs number or the unsuccessful migration of PGCs (Kuo *et al*., 2009). In Nile tilapia, there are three predicted *pum* genes, but their molecular functions have not been studied. Overall, while evidence in vertebrates indicates pivotal roles of *nanos*, *piwil*, *dnd*, *vasa* and *pum* genes on PGCs and germline development, these genes have not been characterised in early ontogenic development of Nile tilapia.

The aim of this study was to profile expression patterns of eleven candidate genes reported to play a central role in PGC specification and development during early development of Nile tilapia. In order to optimise a sterilisation method and minimise pleiotropy and potential side effects associated with PGC ablation, target genes should be exclusively expressed in gonads and have no apparent roles during embryonic development apart from PGC maintenance and survival. The refined target list would enable subsequent research into gene editing techniques that have the potential to disrupt PGC development in this species.

2.2 Materials and Methods
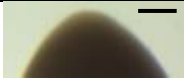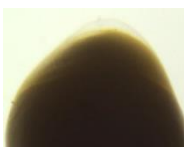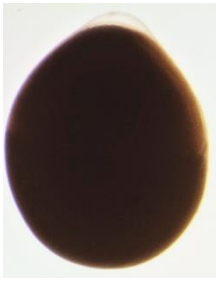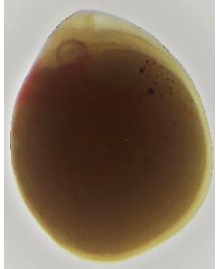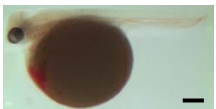
2.2.1 Embryo development of *O. niloticus*

A population of all male (XY) Nile tilapia embryos was produced at the Tropical Aquarium Facility at the Institute of Aquaculture, University of Stirling by crossing YY supermale (Scott *et al*., 1989) with XX female broodstock. 100% male ratio was confirmed in a total of 158 progeny (Chapter 3). Eggs were stripped from a single mature female Nile tilapia into a Petri dish. Before fertilisation, some unfertilised eggs were collected and preserved (as described below) for future analysis, the remainder of the eggs were fertilised with sperm from a single YY supermale. The fertilised eggs were cultured in a tumbling egg system (< 500 embryos in a 700 mL culture chamber) at $27 \pm 1°C$ and 12L:12D photoperiod. At routine intervals, developmental progression was checked using a dissection microscope, and samples of embryos at key developmental stages based on Fujimura and Okada (2007) (Table 2.1) were collected as stored in RNAlater at 4°C until RNA extraction.

2.2.2 RNA extraction and cDNA synthesis

Total RNA from embryo samples, either as pools or as individual samples (Table 2.1), was extracted using the routine TRI reagent (Sigma-Aldrich, St. Louis, USA) extraction method, following mechanical disruption (Mini-BeadBeater, BioSpec Products, Bartlesville, USA) of the samples with 3 mm glass beads (EMD Millipore, Burlington, USA). RNA integrity and yield were confirmed by gel electrophoresis and spectrophotometry, respectively, and then all RNA samples were treated with DNase I (DNA-free DNA Removal kit, Thermo Fisher, Waltham, USA) to remove residual genomic DNA (gDNA) contamination. cDNA was reverse transcribed from 400 ng of DNase I treated RNA (High capacity cDNA reverse transcription kit, Applied Biosystems, Foster City, USA) using a blend of random hexamer & anchored oligo dT primer at a 3:1 ratio. Once produced, cDNA samples were stored at -20°C prior to analysis.

Table 2.1. Sampled developmental stages of *O. niloticus* from unfertilised eggs to hatching stage in chronological order based on Fujimura and Okada (2007). Scale bar = 0.5 mm. hpf: hours post fertilisation. *n*, number of samples anlaysed

Developmental Stage	hpf	Description of development	Figure	Pooled No. of egg/embryo per sample	*n*
Egg	0	Unfertilised eggs		5	5
Zygote	0-1.5	1-cell		10	6
Cleavage	1.5-2	2-cell		10	6
	2	4-cell		10	4
	3	8-cell		10	6
Blastula	4-12	Early blastula		6	6
Gastrula	22-26	Gastrula, epiboly=30-50%		6	6
Segmentation	26-30	Neurula		6	6
Pharyngula	60-72	Onset of blood circulation		2	6
Hatching	90-110	Jaw extension		1	6

2.2.3 Quantitative PCR analysis

The Nile tilapia sequences of *nanos1a, nanos1b, nanos2*, *nanos3*, *piwil1*, *piwil2*, *dnd1*, *vasa*, *pum1*, *pum2*, *pum3*, *elf1α, gapdh* and *ß-actin,* were identified in NCBI and primers (Table 2.2) were designed using Primer-BLAST in NCBI. In the case of *vasa*, there are three *vasa* homologs (accession No. AB649031-AB649033) (Conte *et al*., 2017; Fujimura *et al*., 2011); however, the primer pair designed by Pfennig *et al*. (2012) which was used in the current study amplifies all of three copies. In the case of genes which have multiple potential transcript variants (*piwil1*, *dnd1*, *pum1* and *pum2*) the primer pair was designed on the common sequence and thus would measure all transcript variants expressed. For each target, absolute quantification PCR assays were designed and validated. A standard curve was generated for each target from a serial dilution of a linearised plasmid which had previously had the target-specific partial cDNA fragment cloned within it using pGEM T-easy vector systems (Promega, Madison, USA). The cloned partial sequence within each plasmid was confirmed to have 100 % sequence identity prior to use within the assay by Sanger sequencing (GATC Biotech, Konstanz, Germany). Each qPCR reaction was of a total volume of 10 μL containing 2.5 μL of cDNA (1/10 diluted), 5 μL of SYBR green mix (Luminaris Color HiGreen qPCR master Mix, Thermo Fisher, Waltham, USA), 0.3 μM of each forward and reverse primer and MilliQ (EMD Millipore, Burlington, USA) water up to 10 μL. The qPCR thermal cycling protocol was: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, annealing for 30 sec at temperatures optimised for specific primers (Table 2.2) and 72°C for 30 sec, and completed with melt curve analysis. All samples were analysed in duplicate together with non-template controls and standard samples. The qPCR efficiency of all targets ranged from 0.91 to 1, and the linearity (R^2) of all standard curves ranged from 0.998 to 1. The standard curves of *nanos1a*, *nanos3*, *piwil1* and *piwil2* had an absolute range from 10^2 to 10^7 copies per reaction, and the rest of targets had an absolute range from 10^1 to 10^6 copies per reaction. In all assays, melt curve analysis confirmed a single product was generated and no primer dimer artefacts were evident. The absolute copy numbers of every target gene were calculated based on its standard curve. This was thereafter normalised using the geometric mean of *elf1α*, *gapdh* and *ß-actin*, as this was determined to be the most stable normalisation approach following ranking of expression stability of all possible single & combinations of the three reference genes using Normfinder (http://moma.dk/normfinder-software) (Andersen *et*

al., 2004). To aid interpretation of the results, the pattern of the relative RNA abundance of each target gene during larval ontogeny has been corrected with respect to abundance in the unfertilised eggs, where the abundance at this stage has been given the nominal value of 1.

2.2.4 Tissue distribution

For tissue screening, various tissues (brain, pituitary, eye, heart, intestine, spleen, kidney, liver & gonads) were removed from a male and female adult tilapia following euthanasia using an approved Schedule 1 method (overdose of Benzocaine and severing of spinal column) and the samples were stored in RNAlater at 4°C prior to RNA extraction (the tissue samples all derived from a single male except for ovarian sample which came from a single female tilapia from the same stock). Total RNA was extracted and cDNA synthesised using the same method described previously. The tissue-specific abundance of the chosen targets was assessed by routine PCR where each reaction consisted a total volume of 5 μL containing 2.5 μL MyTaq mix (Bioline, London, UK), 0.5 μL of cDNA (1/10 diluted, 2.5 ng), 0.1 μM of each forward and reverse primer and MilliQ (EMD Millipore, Burlington, USA) water up to 5 μL. The PCR thermal cycling protocol was as follows: 95°C for 1 min, followed by 32 cycles (except for *ß-actin* of 28 cycles) of 95°C for 15 sec, annealing for 20 sec at temperatures optimised for specific primers (Table 2.2) and 72°C for 30 sec and final elongation at 72°C for 2 min. The presence of each gene target was assessed by gel electrophoresis.

Table 2.2. Primers used for qPCR and tissue screening PCR

Gene	Accession No.	Forward primer (5'-3')	Reverse primer (5'-3')	Annealing temp (°C)	Analysis
nanos1a	XM_003447766.4	TCTCAGGCCATACGAACAC CTCG	CTCTGAGCCTGTTTGCGT CTTCG	62	qPCR
		TAAACTGTGCGGCTCTGAC AGGA	CTCTGAGCCTGTTTGCGT CTTCG	62	Tissue screen
nanos1b	XM_005467222.3	GAGCCCCTTCCAAAATCAG CTCG	CATGCCGAACAGATCGAA CCCAG	62	qPCR
		GAGCCCCTTCCAAAATCAG CTCG	GCTGGGATGGCTGGTCTT TTGAC	62	Tissue screen
nanos2	XM_005448855.3	CGGGAAAGTTTTCTGCCCC ATCC	AGAACTTGGCCCCTGTCT CCATC	62	qPCR
		GTTCGAGAGACTGTGCAGC AACC	AGAACTTGGCCCCTGTCT CCATC	62	Tissue screen
nanos3	XM_005460553.3	GGAGTGTGACATGAGCCGA GCTA	AACTCGTTAGTGCACATT CGCGG	62	qPCR
		GGAGTGTGACATGAGCCGA GCTA	CATACTGCCGCAGAAAGG GACAC	62	Tissue screen
piwil1	XM_003445546.3	ATGATCGTGGGCATCGACT GCTA	ACCACCTGCTCATGCTTT GGTTG	62	qPCR
		ATGATCGTGGGCATCGACT GCTA	CGGGTGACCTCTGAGTCA ACGAT	62	Tissue screen
piwil2	XM_003445662.3	TGCCATCAAGAAGCTGTGC TGTG	CTGGGAAATTGTGCGGAC GTTGA	62	qPCR
		CAAGAAGCCACCGACCTTT CCAC	CTGGGAAATTGTGCGGAC GTTGA	62	Tissue screen
dnd1	XM_003454288.4	CACGGGACACGTATGAGGA CATC	ATATTTGGCATACGCAAA GCCGC	60	qPCR
*vasa**	AB032467.1	CGATGAGATCTTGGTGGAT G	CATGAGATCCCTGCCAGC AGA	58	qPCR
pum1	XM_013270654.2	GCTAACTGGTAAGAAGTTC TGGGAA	CGGGACACCATGATTGGC TG	60	qPCR
pum2	XM_003446310.4	GCGACATTGTCCGTGAGCT TGAT	CAAAAACCTGTCCCTGGA AGGCA	60	qPCR
pum3	XM_005470475.3	TCACAGATGAGCTGTACGG CAAC	CATGGGTGTGAGAACCTG CTTCA	60	qPCR
elf1α	NM_001279647.1	CTGGACAAACTGAAGGCTG AGCG	AAGTCTCTGTGTCCAGGG GCATC	62	qPCR
gapdh	NM_001279552.1	TGGATCTGACATGCCGTCT CTCC	AATCAGAGGACACCACCT GGTCC	62	qPCR
β-actin	KJ126772.1	ATCTACGAGGGTTATGCCC TGCC	CTGTGGTGGTGAAGGAGT AGCCA	62	qPCR
		ATCTACGAGGGTTATGCCC TGCC	CGATGCCAGGGTACATGG TGGTA	60	Tissue screen

* based on Pfennig *et al*. (2012).

2.2.5 Statistics

Statistical analysis was performed using Minitab 17 (Minitab Inc., State College, USA). The relative expression values were $\log_{10}$-transformed (except for *pum1*) after which they met the assumptions of normality and homogeneity of variance. Then, significant differences between different developmental stages were tested by one-way ANOVA, followed by Tukey's HSD test ($p < 0.05$). Data are presented as mean $\pm$ standard deviation (SD).

2.3 Results

2.3.1 Absolute transcripts number in unfertilised eggs

In unfertilised eggs, the maternally transferred RNAs were detectable in all but one target (Table 2.3). *piwil1* had the greatest abundance being recorded at 53,524 $\pm$ 3,090 copies per ng total RNA. The abundance of *nanos3*, *piwil2, vasa* and *pum2* was an order of magnitude lower, ranging between 1,119 $\pm$ 257 and 9,293 $\pm$ 374 copies per ng total RNA. The abundance of *nanos1a*, *dnd1* and *pum3* was a further order of magnitude lower, registering between 181 $\pm$ 17 and 605 $\pm$ 227 copies per ng total RNA, while the abundance of *nanos1b* and *pum1* was an order of magnitude lower still, and therefore approaching the detection limits of the technique with *nanos2* being undetectable in unfertilised eggs.

Table 2.3. Absolute copy numbers of maternal transcripts in unfertilised eggs ($n = 5$). Data are shown as mean ± SD

	Absolute copy number per ng total RNA
nanos1a	181 ± 17
nanos1b	5 ± 1
nanos2	below detection range
nanos3	9,293 ± 374
piwil1	53,524 ± 3,090
piwil2	4,614 ± 202
dnd1	605 ± 227
vasa	1,851 ± 93
pum1	15 ± 5
pum2	1,119 ± 257
pum3	290 ± 59

2.3.2 Quantitative relative transcript levels during ontogenic development

All targets were detectable throughout larval ontogeny with the exception of *nanos2* which was undetectable until hatching at which point abundance was at the detection limits of the methodology (data not shown). There were three apparent patterns of target abundance observed within the dataset.

The largest group, consisting of *nanos3, piwil1, piwil2, dnd1* and *vasa,* could be characterised as maintaining abundance until early blastula stage, after which the target abundance significantly reduced until hatching stage with the abundance steadily decreasing by four or five (*vasa* only) orders of magnitude over this developmental window (Fig. 2.1C, D, E, F&G).
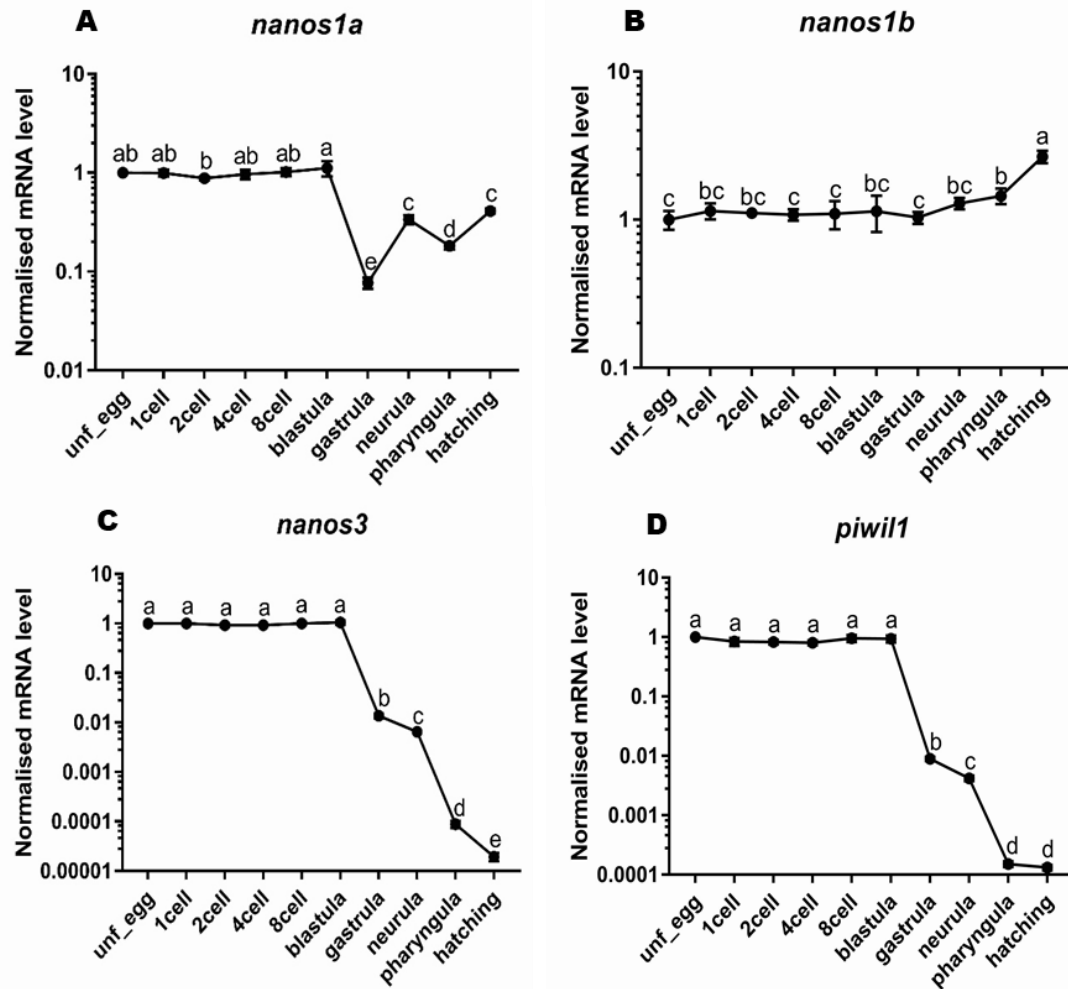
Figure 2.1. Ontogenic expression patterns. Normalised relative mRNA transcript level of (A) *nanos1a,* (B) *nanos1b,* (C) *nanos3,* (D) *piwil1,* (E) *piwil2,* (F) *dnd1,* (G), *vasa,* (H) *pum1,* (I) *pum2* and (J) *pum3*. Data are presented as mean ± SD ($n = 5 - 6$). Y-axis is shown in logarithmic scale. Superscripts (a-f) denote statistically significant difference between developmental stages. Unf_egg: unfertilised eggs. (continued on next page)
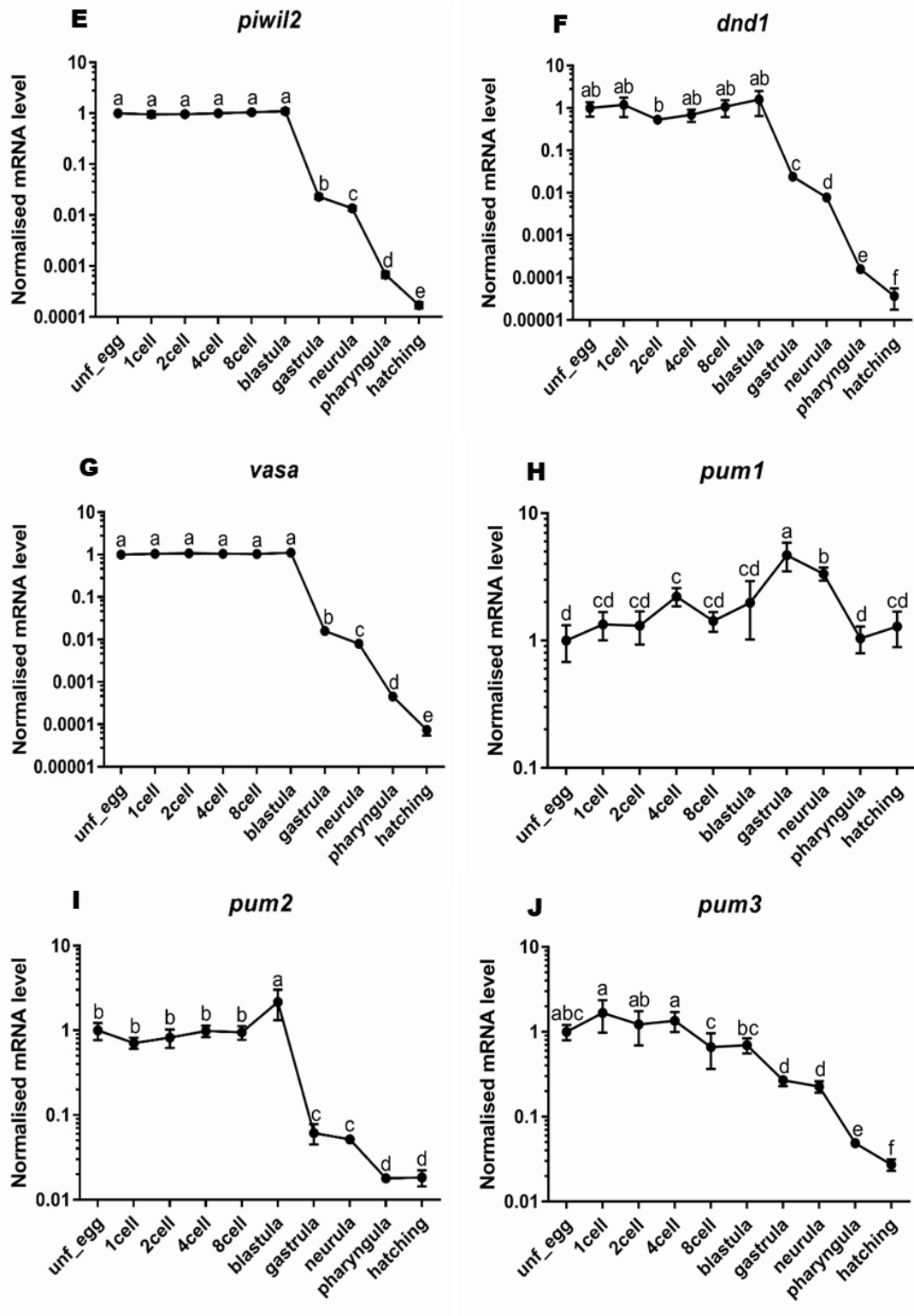
Figure 2.1. (continued from previous page).

The second group consisted of *pum2* and *pum3* and showed similarities to the first group in that they could be characterised as having relatively stable abundance until blastula stage; thereafter a significant reduction (with respect to unfertilised eggs) in abundance was recorded. This reduction in transcript abundance continued until hatching stage; however, unlike the first group, this reduction was in the region of two orders of magnitude in comparison to levels recorded in unfertilised eggs (Fig. 2.1I&J).

In the final group, consisting of *nanos1a, nanos1b* and *pum1,* the abundance remained relatively constant over the duration of the study, recording no absolute changes more than an order of magnitude different from unfertilised eggs (Fig. 2.1A, B&H). Instead of target degradation in the later developmental stages, all targets showed evidence of active expression with the abundance in *nanos1a* being significantly elevated from neurula stage onwards with respect to the gastrula stage. Equally, for *nanos1b,* the expression levels increased significantly from pharyngula to peak at hatching, while in *pum1* there was a significant elevation in the expression at gastrula and neurula stages before the expression levels returned to being comparable to the abundance in unfertilised eggs. (Fig. 2.1A, B&H).

2.3.3 Tissue distribution

In adult tilapia *nanos1a*, *nanos1b, pum1, pum2* and *pum3* were expressed in a wide range of tissues including brain, pituitary, testis and ovary (Fig. 2.2). *nanos2* was exclusively expressed in testis, and *nanos3*, *piwil1*, *piwil2, dnd1* and *vasa* were gonad-specific, being expressed in both testis and ovary (Fig. 2.2).
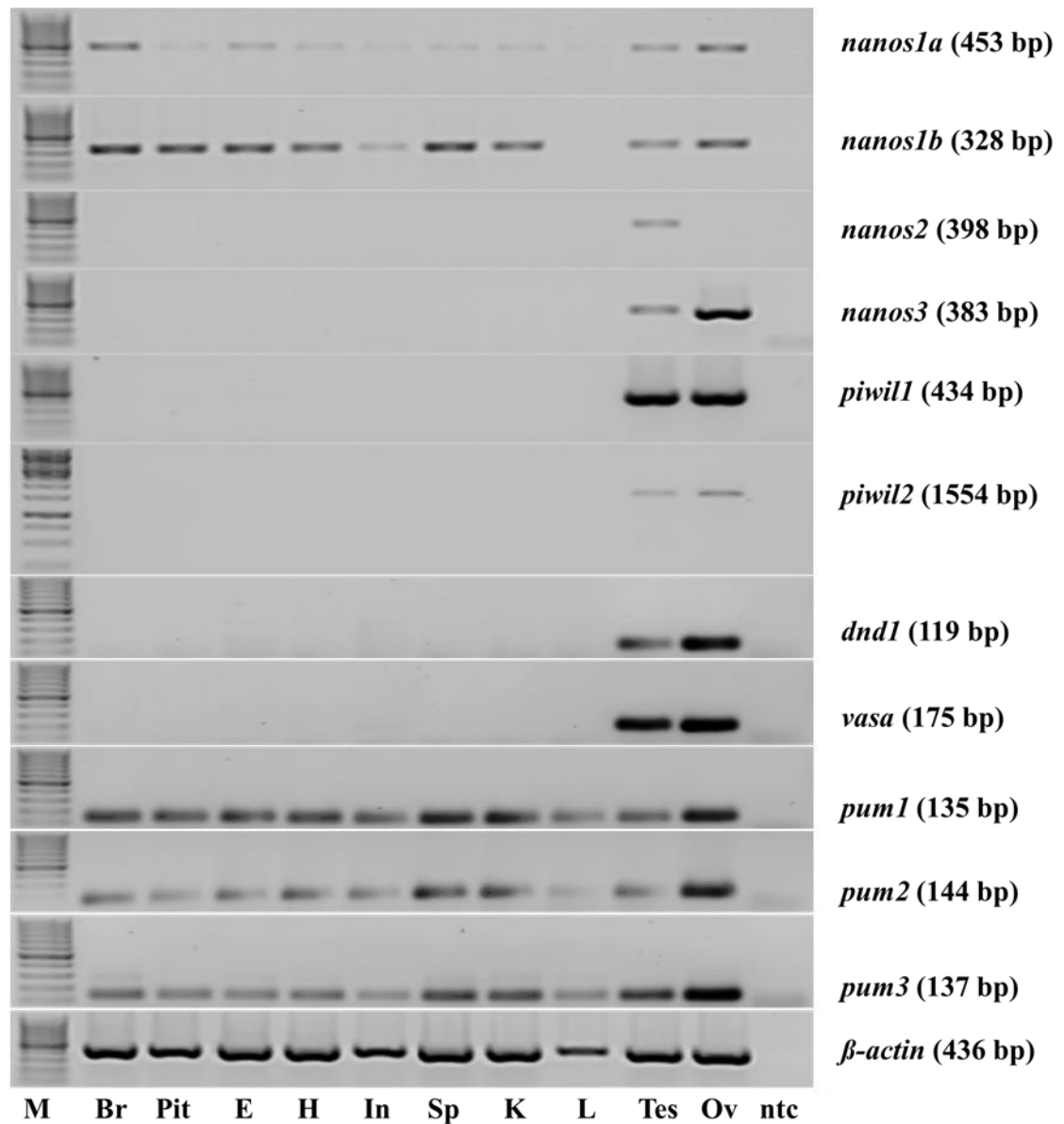
Figure 2.2. Tissue distribution of *nanos*, *piwil, dnd1, vasa* and *pum* genes in Nile tilapia. Br, brain; Pit, pituitary; E, eye; H, heart; In, intestine; Sp, spleen; K, Kidney; L, liver; Tes, testis; Ov, ovary. Marker (M) is 100 bp ladder except for *piwil2* which is 1 kb. Male, 24.1 cm in body length (BL), 269.1 g in body weight (BW), 0.22% gonadal somatic index (GSI); female, 19.5 cm in BL, 124.4 g in BW, 0.4% GSI.

2.4 Discussion

In this study, 11 candidate genes were identified as putative targets for gene knockout to induce sterility in Nile tilapia, based on their published regulatory role in germ cells to repress somatic cell fate in a range of vertebrate species. The optimal candidate gene target should be clearly gonad-specific and also not associated with the normal developmental process during embryogenesis. The targeted removal/silencing of such a candidate gene should, therefore, specifically remove germ cells without having effects on the early developmental process. Maternal deposition of the transcripts would also be desirable as this would imply that the target gene plays a potential role in PGC specification and/or maintenance. With this in mind, an ontogenic expression study with tissue screening was undertaken to rationalise the target gene list further in Nile tilapia.

Through the observation of the ontogenic pattern of the relative RNA abundance of each candidate gene, inappropriate targets could be excluded based on the presence or absence of apparent zygotic expression. The active zygotic expression would imply a role played by the gene in normal development rather than specification and maintenance of PGCs. This was shown in Atlantic salmon and medaka where maternally provided germ cell-specific genes such as *dnd, tdrd1*, *tdrd7*, *tdrd7-2*, *dazl, dazl-2*, *piwil1* showed a rapid decrease during the maternal to zygotic transition (MZT) without significant increase in mRNA levels (Kleppe *et al.*, 2015; Liu *et al.*, 2009; Xu *et al.*, 2007; Zhao *et al.*, 2012). The MZT has been reported to occur from fertilisation when maternal transcripts begin to be degraded, and generally, a bulk zygotic transcription initiates at midblastula transition (MBT) when cell cycles become longer as described in a range of both vertebrate and invertebrate species (reviewed in Langley *et al.*, 2014). In the current study, the degradation of the majority of the maternal transcripts of *nanos3, piwil1*, *piwil2, dnd1* and *vasa* occurred between the blastula and gastrula stages during MZT and there was no apparent zygotic transcription to recover the abundance levels of these targets up to the point of the hatch when the study ended. The functional mechanism of this degradation has not been studied as yet in tilapia and requires further research. Evidence from other species including zebrafish, *Xenopus* and *Drosophila* suggested that this degradation is controlled by a family of microRNAs including zebrafish miR-430, *Xenopus* miR-427 and *Drosophila* miR-309, known to promote deadenylation and decay of maternal mRNAs during the MZT (Bushati *et al.*, 2008; Giraldez, 2010; Giraldez *et al.*, 2006; Lund *et al.*, 2009). In contrast to this group of genes, *nanos1a*, *nanos1b, pum1*

and *pum2* all showed a significant increase at different developmental stages, while *pum3* did not show a sudden decrease in its transcripts level, which all imply an active zygotic expression. The ontogenic expression patterns of these genes suggest they play an active role in normal embryonic development apart from PGCs in Nile tilapia and thus should be excluded from the potential target list for sterilisation through gene editing techniques (Kleppe *et al*., 2015; Liu *et al*., 2009).

Gonad-specific expression is critically important to ensure that gene editing has no off-target effects. In the current study, both *nanos1a*, *nanos1b* as well as *pum1*, *pum2* and *pum3* were all ubiquitously expressed in a range of tissues. Such a wide-ranging expression of *nanos1a* and *nanos1b* has been reported in Chinese sturgeon (*Acipenser sinensis*) (Ye *et al*., 2012). However, contrasting findings have also been reported for tissue-specific expression of *nanos* genes with differential specificity in tissue expression in medaka (Aoki *et al*. 2009) and lack of gonadal expression in orange-spotted grouper (*Epinephelus coioides*) (Sun *et al*. 2017). A wide tissue distribution was also reported for *pum1*, *pum2* and *pum3* in zebrafish, while expression of another *pumillio* gene, *puf-A,* was found in various tissues and the knockdown of this gene resulted in eye defects, reduction of PGCs number and mis-migration of PGCs, indicating its important role in the development of eyes and survival and migration of PGCs (Kuo *et al*., 2009). Therefore, it can be concluded that all of these targets are likely to be involved in various biological roles during the embryonic development of Nile tilapia and not only in gonadal development. In contrast, *nanos2, nanos3, piwil1, piwil2, dnd1* and *vasa* all appeared to be gonad-specific in Nile tilapia. *nanos2* was only expressed in testis in this study as shown in mice (Tsuda *et al*., 2003); however, in many teleosts species including zebrafish, medaka and rainbow trout, the expression of *nanos2* was detected in both oogonia and spermatogonia in adult gonads (Aoki *et al*., 2009; Beer and Draper, 2013; Bellaiche *et al*., 2014). While *nanos3* was observed in both testis and ovary in the current study examined by RT-PCR, previous research reported *nanos3* to be specific to ovarian tissue in tilapia using both RT-PCR and *in situ* hybridisation (Li *et al*., 2014). These contrasting results may be due to a temporal variation of this gene expression between the two studies and therefore further research is needed. In other teleosts, while *nanos3* remains a gonad-specific gene, the gender specificity is not consistent with it being reported in both the testis and ovary of rainbow trout (Bellaiche *et al*., 2014), in contrast to being ovary-specific in zebrafish (Köprunner *et al*., 2001; Beer and Draper, 2013), medaka (Herpin *et al*., 2007) and grouper (Sun *et al*., 2017). Taking an advantage of ovarian germ cell

specificity of *nanos3* in zebrafish, Zhou *et al*. (2018) produced transgenic fish in which transgene expression is driven by *nanos3* promoter. In the current study, both the *piwil* homologs were present in both gonadal tissues: this is in agreement with the published studies in zebrafish, medaka and Atlantic salmon for *piwil1* (Houwing *et al*., 2007; Kleppe *et al*., 2015; Zhao *et al*., 2012) and Atlantic salmon and rainbow trout for *piwil2* (Kleppe *et al*., 2015; Yano *et al*., 2008). However, tissue-specificity is also not consistent across teleosts as the expression of *piwil2* in medaka was reported in a wide range of tissues (*e.g.* brain, gill, heart, liver, kidney, spleen, intestine, ovary and testes) (Zhao *et al*., 2012). The gonad specificity of *vasa* observed in the current study is in agreement with the previous work in zebrafish and Atlantic cod (Presslauer *et al*., 2012; Yoon *et al*., 1997), while the observed gonad specificity of *dnd1* aligns with the work in a wider range of teleosts species including zebrafish, medaka, Atlantic salmon and turbot (*Scophthalmus maximus*) (Kleppe *et al*., 2015; Lin *et al*., 2013; Liu *et al*., 2009; Weidinger *et al*., 2003). When considered together these results demonstrate the need to confirm tissue-specificity in target species. Evidently, the diversity in reproductive strategy and development, as well as sex determination and differentiation processes in teleosts, makes it inappropriate to assume comparable functional roles for these genes (Kobayashi *et al*., 2013; Li *et al*., 2011). Therefore, in terms of prioritising targets to be considered for gene editing to induce sterility in Nile tilapia, *nanos1a*, *nanos1b*, *pum1*, *pum2* and *pum3* should clearly be excluded on the grounds of lack of gonad specificity. Arguably, *nanos2* should also be omitted until its functional role in PGC specification and/or early gonadal development, prior to differentiation, is confirmed.

GP components are essential to maintaining germ cell fate, otherwise, they become somatic cells (Cinalli *et al*., 2008; Lai and King, 2013; Tada *et al*., 2012). *Nanos*, *pum*, *vasa, dnd* and *piwil* genes are known to be involved in translational repression in germ cells, which is thought to be key to the maintenance of the germline (Asaoka-Taguchi *et al*., 1999; Carmell *et al*., 2007; Cox *et al*., 2000; Lai and King, 2013; Mickoleit *et al*., 2011). In Nile tilapia, RNA transcripts of all the investigated genes were present in the unfertilised eggs demonstrating maternal transfer, with the exception of *nanos2*. The transcript level of *nanos2* was below the detection limits until hatching stage which suggests that it is unlikely to be involved in the specification of PGCs. Among the maternally deposited RNAs in this study, *nanos3, piwil1, piwil2*, *dnd1* and *vasa* did not show evidence of significant zygotic expression, as discussed previously. Such a degradation in abundance is a common feature of many maternal transcripts which are

components of GP such as *nanos3, vasa, dazl, dnd* and *tdrd7* in teleosts (Knaut *et al*., 2000; Köprunner *et al*., 2001; Kosaka *et al*., 2007; Strasser *et al*., 2008; Weidinger *et al*., 2003; Xu *et al*., 2005). It should, however, be acknowledged that PGC-specific expression of *nanos3* in embryos was reported in various teleost species including zebrafish (Beer and Draper, 2013; Köprunner *et al*., 2001), medaka (Herpin *et al*., 2007), Atlantic cod (Presslauer *et al*., 2012) and Atlantic salmon (Škugor *et al*., 2014a). This suggests that discrete PGC expression may have not been captured by the analysis of whole embryo expression which recorded a decrease by four or five orders of magnitude of these genes. Therefore, we can postulate that in Nile tilapia *nanos3*, *piwil1*, *piwil2*, *dnd1* and *vasa* are maternally deposited GP components based on their ontogenic pattern of RNA abundance.

When considered together, the current results suggest that of the eleven candidate genes proposed, five (*nanos3*, *piwil1*, *piwil2*, *dnd1* and *vasa*) would be viable targets for gene editing to induce sterility in tilapia. To date, only *nanos* genes have been targeted for this purpose in Nile tilapia. In *nanos3* KO XX tilapia using CRISPR/Cas9, Li *et al*. (2014) reported the loss of PGCs at hatching stage and germ cell-deficient gonads at 60 dph, suggesting that maternal transcripts of *nanos3* are not enough to maintain germ cells at the hatching stage, andzygotic transcripts during ontogenic development are vital for the survival of PGCs. Furthermore, knockdown of *nanos3* using morpholinos revealed its essential role in PGC specification in zebrafish (Köprunner *et al*., 2001) and a similar key role in PGCs and germ cell development has been demonstrated in *Drosophila* (Forbes and Lehmann, 1998). Interestingly, Li *et al*. (2014) also targeted *nanos2* by CRISPR/Cas9 in Nile tilapia, which resulted in the loss of PGCs at hatching stage and germ cell-deficient gonads at 90 dph. This is in agreement with mammalian research where *nanos2* KO mice had no germ cells (Tsuda *et al*., 2003). From the current study, *nanos2* has been excluded from the candidate list based on the screening results and primarily the apparent testis-specific expression. The successful disruption of PGC development via *nanos2* KO suggests that specificity of expression in differentiated gonads is not necessarily an accurate reflection of the role of the gene in early PGC development which warrants further investigation. Of the remaining candidates (*piwil1, piwil2*, *dnd1* and *vasa*) there is no published research exploring their functional role in PGC and ultimately gonadal development in tilapia. Evidence from other teleosts is indicative that these warrant further research. For example, in zebrafish *ziwi* and *zili* mutants, the germ cells are significantly reduced or not present (Houwing *et al*. 2007 &

2008), while morpholino knockdown of *piwil1* resulted in a reduction in the number of PGCs and defects in PGC migration in medaka (Li *et al.*, 2012). With respect to *dnd1*, Gross-Thebing *et al.* (2017) reported that in *dnd*-deficient zebrafish the PGCs transdifferentiate into somatic cells, equally *dnd* KO using CRISPR/Cas9 in Atlantic salmon resulted in germ cell loss (Wargelius *et al.*, 2016). In contrast, Škugor *et al.* (2014b) reported that the knockdown of *dnd* in Atlantic cod caused a decrease of expression levels of *vasa*, *nanos3* and *tdrd7*, but also showed evident pleiotropic effects suggesting, in Atlantic cod at least, that *dnd* is involved in various developmental processes including suppression of reproduction. These results again clearly demonstrate the need for species-specific research to avoid deleterious effects on the development and health of individuals as part of gene editing research. Finally, in relation to *vasa*, it has been widely reported that the mutation of *vasa* leads to the loss of PGCs, defective germ cell proliferation and differentiation in the gonad across animal phyla, including fruit flies, nematode worms, frogs and mice (Hay *et al.*, 1988; Ikenishi and Tanaka, 1997; Gruidl *et al.*, 1996; Shinomiya *et al.*, 2000). In zebrafish, however, *vasa* knockdown does not appear to affect the specification of PGCs, suggesting the maternal and zygotic *vasa* transcripts during ontogenic development are not essential (Braat *et al.*, 2001), but *vasa* KO zebrafish revealed that the zygotic transcript of *vasa* is required for germ cell differentiation and survival in the later stage (Hartung *et al.*, 2014). In medaka, knockdown of *vasa* led to a number of PGCs not entering the putative gonad and the ectopic PGCs that lack *vasa* survived and even increased their number, suggesting that *vasa* may play an important role in apoptosis of ectopic PGCs as well as the successful migration of PGCs (Li *et al.*, 2009).

This research has helped rationalise the candidate genes to be taken forward into functional disruption studies with the aim to develop interventions that will induce sterility in Nile tilapia by disrupting normal PGC development. There are nearly 4 million tonnes of Nile tilapia being farmed annually across the global (FAO, 2017); however, production productivity is reduced due to losses associated with maturation in culture. While the industry addresses this challenge currently through the farming of single sex (all male) stocks (Phelps and Popma, 2000), there is a demand for development of alternative sterilisation methods. To this end, further research will be conducted using genome-editing techniques to analyse the functions of *piwil* genes in PGCs, which were prioritised by this study, with an aim to induce sterility in Nile tilapia.

CHAPTER 3

IMPACT OF HIGH TEMPERATURE ON REPRODUCTIVE COMPETENCE IN MALE NILE TILAPIA

3.1 Introduction

The disruption of primordial germ cells (PGCs) at early developmental stages has been proposed as a possible means to induce sterility in farmed fish (Wong and Zohar, 2015a). To rationalise the choice of candidate genes for knockout using CRISPR/Cas9 genome editing 11 candidate genes were screened to elucidate their association with the specification and maintenance of PGCs during the ontogenic development in Nile Tilapia (Jin *et al*., 2019). To build on this, the work of Nakamura *et al*. (2015) and Pandit *et al*. (2015) opened up an interesting research paradigm where the functional role of these candidates in relation to temperature-induced germ cell loss can be further explored.

Water temperature is a critical environmental factor for fish gonadal development and spawning. High temperature treatment can induce sex reversal or suppress reproduction in fish. In genetically female zebrafish exposed to 37°C between 15 - 25 dah, oocytes at early diplotene stage regressed and masculinised (Uchida *et al*., 2004). In medaka, high temperature (33°C) inhibited proliferation of germ cells and induced female-to-male sex reversal associated with an increase in plasma cortisol (Hayashi *et al*., 2010). Baroiller *et al*. (2009) suggested apoptosis and/or proliferation of PGCs caused by high temperature during the labile period of sex differentiation could be a critical factor for sex reversal in fish. In addition, water temperature above the natural range for the species can hinder gonadal development and cause germ cell loss in teleosts (de Alvarenga and de França, 2009; Ito *et al*., 2003; Soria *et al*., 2008; Strüssmann *et al*., 1998; Uchida *et al*., 2004). Germ cell death caused by heat treatment was reported in male pejerrey, showing nuclear pyknosis or eosinophilia (Ito *et al*., 2003; Soria *et al*., 2008) and in fugu (Lee *et al*., 2009). Sterility was reported in female Patagonian freshwater pejerry and pejerrey following exposure to 27 – 28.5 and 29°C, respectively for 112 – 135 days from 1 – 5 weeks after hatching (Strüssmann *et al*., 1998). All together, these data suggest species-specific thermal sensitivity of germ cell survival.

It has been reported that a high temperature treatment at 37°C for 40 – 50 days from 3 dah can induce complete sterility in female Nile tilapia (Chitralada strain) and male Mozambique tilapia (Nakamura *et al*., 2015; Pandit *et al*. 2015). Adult male Nile tilapia exposed to 30 - 35°C for a week showed increased germ cell differentiation with lower Sertoli and Leydig cell proliferation compared to fish exposed to 20 - 25°C, but the number of germ cells was not affected (de Alvarenga and de França, 2009; Lacerda *et al*., 2006). While several studies have shown the temperature-specific effects of heat

treatments on gonadal germ cells and overall suppression of gametogenesis in fish, the underlying functional mechanisms leading to germ cell death remains unclear.

Studies in fish have shown that the specification of PGCs is regulated by maternally provided GP (Raz, 2003). The maternally deposited GP consists of maternal mRNAs and proteins of germ cell-specific genes including *piwil* genes and *nanos3* (referred to as *nanos1* in zebrafish in previous literature) which are found in PGCs and shown to have essential roles in PGC specification, germ cell development and/or maintenance, but their functions are unclear in many teleost species (Draper *et al*., 2007; Houwing *et al*., 2007; Köprunner *et al*., 2001). *Piwil1* and *piwil2* are germline specific argonautes known to maintain germline fate and suppress transposon activity together with piRNA (Houwing *et al*., 2007, 2008). In Nile tilapia, *piwil1* and *piwil2* transcripts are maternally deposited in eggs and exclusively expressed in both ovary and testis in Nile tilapia (Jin *et al*., 2019). In this species, both Piwil1 and Piwil2 proteins were present in various stages of testicular germ cells from spermatogonia to spermatid but not in spermatozoa (Xiao *et al*., 2013). *nanos2* and *nanos3* are RNA-binding zinc finger proteins and *nanos2* is known to be a key regulator for the maintenance and modulation of SSCs self-renewal (Hofmann, 2008; Sada *et al*., 2009; Suzuki *et al*., 2009). In Nile tilapia, *nanos2* is expressed exclusively in testis and transcripts appeared to not be maternally provided in eggs (Jin *et al*., 2019). *nanos2* has been reported as a marker for putative SSC in rainbow trout and Nile tilapia (Bellaiche *et al*., 2014; Lacerda *et al*., 2013). On the other hand, *nanos3* was reported to play an important role in PCG survival and migration (Doitsidou *et al*., 2002; Draper *et al*., 2007). The expression of *nanos3* was reported in undifferentiated spermatogonia and the early stages of differentiating spermatogenic germ cells in both mammals (Suzuki *et al*., 2009) and rainbow trout (Bellaiche *et al*., 2014). In Nile tilapia, *nanos3* is maternally transferred to the zygote and expressed in testis as well as in ovary (Jin *et al*., 2019). Glial cell line-derived neurotrophic factor (GDNF) family receptor alpha-1 (*gfra1*) was also suggested as a potential marker of SSCs, as it is expressed exclusively in single type A undifferentiated spermatogonia in Nile tilapia (Lacerda *et al*., 2013). Therefore, *piwil1, piwil2, nanos2*, *nanos3* and *gfra1* are considered as suitable markers for testicular germ cells at different stages.

Excessive apoptosis in testis can cause infertility in humans (Agarwal and Said, 2005; Tesarik *et al*., 1998). As reviewed by AnvariFar *et al*. (2017) heat stress is a major environmental stressor and it can cause germ cell loss through the stimulation of apoptotic pathways in fish. Signalling pathways mediating apoptosis involve various

molecules including Bcl-2 family proteins, caspases, cytochrome c, p53, Fas and FADD. Bcl-2 family genes can be divided into two groups based on their role, proapoptotic (*bax, bim, box*) and prosurvival (*bcl-xL, bcl-2, mcl-1*) (Cory and Adams, 2002). The apoptotic pathways are conserved in fish and mammals although some differences can be found such as the lack of a C-terminal region in the FADD in teleosts (AnvariFar *et al*., 2017). While these pathways are well characterised in mammals, very little is known in fish. In mice, rat and cynomolgus monkeys (*Macaca fascicularis*), exposure to elevated temperature caused stage-specific apoptosis of germ cells in testes (Lue *et al*., 1999, 2002; Yin *et al*., 1997). In subadult pejerrey, the exposure to high temperature induces apoptosis in both somatic and germ cells with significant increases of caspase-3 activity in both genders (Ito *et al*., 2008). The severity of heat-induced apoptosis was proportional to the magnitude of the thermal stress and males tended to be more sensitive to thermal stress than females in pejerrey (Ito *et al*., 2008). Still, the understanding of the underlying mechanism of heat-induced apoptosis in gonads is limited in teleosts. Oltval *et al*. (1993) suggested that the balance between proapoptotic and prosurvival Bcl-2 family proteins can determine survival or death of cells, as demonstrated in mouse PGCs where the ratio of Bcl-x and Bax regulate the survival of PGCs and apoptosis (Rucker *et al*., 2000). In teleosts, however, it is unknown which Bcl-2 family members are involved in temperature-mediated germ cell apoptosis.

The high temperature-induced germ cell loss is likely to be associated with a dysregulation of sex steroid hormones as these play essential roles in the survival of germ cells (Billig *et al*., 1996; Lue *et al*., 1999). High temperature disrupts the GnRH-GTHs (FSH and LH) system in fish (Soria *et al*., 2008), resulting in a decrease in the production of sex hormones from gonado-somatic cells. Subsequently, the germ cells die, but the mechanism underlying the link between sex hormones and germ cell survival remains unclear. 11-ketotestosterone (11-KT) is the most abundant sex steroid in male teleosts plasma, whereas testosterone (T) is the major circulating androgen in females (Borg, 1994; Lokman *et al*., 2002; Norris and Lopez, 2011). In support of the functional association, Pandit *et al*. (2015) demonstrated that high temperature induced germ cell loss also resulted in significant reductions in 11-KT, T and estradiol-17β (E2) in female Nile tilapia. On the other hand, heat treatment induced sterile male Mozambique tilapia had normally developed Leydig cells and the level of sex steroid hormones (11-KT, T and E2) was comparable to the control fish (Nakamura *et al*., 2015).

This study aimed to build on the work of Nakamura *et al*. (2015) and Pandit *et al*. (2015) and for the first time investigate, at the molecular level, the impact of high temperature induced germ cell loss in Nile tilapia. To do so, tilapia at 120 dah were chosen to be large enough to obtain testes and blood samples for the subsequent analyses. In this work, key genes associated with both survival and apoptosis of germ cells were profiled to investigate their potential role in associating with germ cell loss. Such a rationalisation of candidates will help prioritise future studies to develop a new means of sterilisation in fish. To this end, the impact of high temperature (36 - 37°C) on the survival of germ cells and reproductive development was investigated in male Nile tilapia. The impact of the high temperature exposure on the germ cell survival and apoptosis was analysed through expression levels of SSC and/or spermatogonia markers (*nanos2, nanos3, piwil1, piwil2* and *gfra1*) and apoptosis-related genes (Bcl-2 family genes, *bcl-xLa, bcl-xLb, baxa, baxl, bcl-2* and *mcl-1a*). In addition, the treatment effects on male reproductive development was analysed through serum sex steroid levels, testis histology and sperm quality including sperm density, motility and fertilisation rate.

3.2 Materials and Methods

3.2.1 High temperature treatment and sampling

All male (XY) progeny were produced by crossing YY supermale (Scott *et al*., 1989) with XX female. A total of 158 progeny were sexed throughout the experiment and all were males. Fish were reared at 27 ± 1°C and a photoperiod of 12L:12D. All male juvenile Nile tilapia at 120 dah were divided into three temperature groups: two high temperature treatments (HT), 35.9 ± 0.1°C (HT1), 36.9 ± 0.1°C (HT2) and a control temperature treatment of 26.7 ± 0.4°C (CT). Water temperature was recorded every 5 min by underwater sensor and a data logger (HOBO Pendant® Onset Computer Corporation, Bourne, USA) and pH (between 6.4 and 7.4) and ammonium level ($<$ 0.1 µg/L of unionized ammonia, NH_3) were monitored on a daily basis during HT period. Timing of samplings was standardised between temperature treatments at 1200, 2100, 3000, 3800 and 5200 degree-days (DD) (Fig. 3.1). Each fish was killed by an approved Schedule 1 method (overdose of Benzocaine and severing of spinal column) at 1200, 2100 and 3000 DD during the temperature treatments and at 3800 and 5200 DD during

the recovery period at 27°C. Body weight, visceral somatic index (VSI), hepatosomatic index (HSI) and gonadal somatic index (GSI) were recorded. Testis samples were either fixed in Bouin's solution for histology (H&E stain) or stored in RNAlater at 4°C prior to RNA extraction for qPCR analysis, respectively (1 lobe for each). All working procedures were carried out in accordance with the United Kingdom Animals (Scientific Procedures) Act 1986 and were approved by the Animal Welfare and Ethical Review Body (AWERB) ethics committee of the University of Stirling.
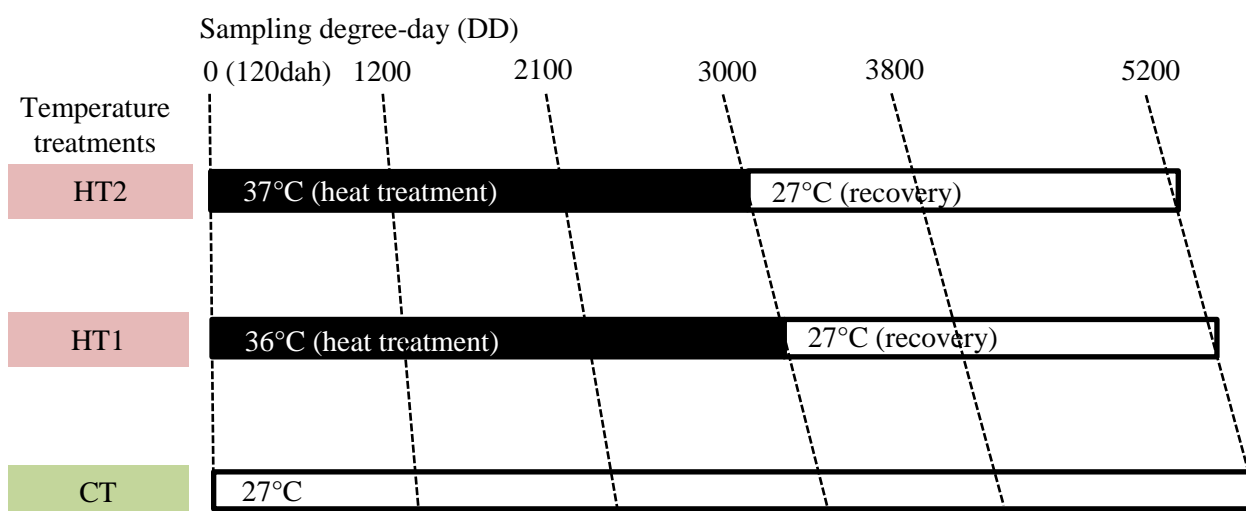
Figure 3.1. Diagram representing the experimental design with three temperature treatments, showing sampling points based on degree-day (DD) for different temperature groups. 0 DD began at 120 dah. The black bar indicates heat treatment period and white bar shows rearing at control temperature of 27°C.

3.2.2 RNA extraction and cDNA synthesis

Total RNA from testis samples ($n = 6$ fish/treatment/time point) were extracted using TRI reagent (Sigma-Aldrich, St. Louis, USA), with the tissue being homogenised prior to extraction using a bead-beater (BioSpec Products, Bartlesville, USA). Total RNA quality was provisionally checked by spectrophotometry (A260/A280 ratio was >1.8) and the RNA integrities were confirmed by checking for the presence of clearly defined ribosomal RNA bands on 1% agarose gel. The RNAs were treated with DNase I to remove gDNAs using DNA-free DNA Removal kit (Thermo Fisher, Waltham, USA). Then, 400 ng of DNase I treated RNAs were used for cDNA synthesis using High capacity cDNA reverse transcription kits (Applied Biosystems, Foster City, USA) with

a blend of random hexamer & anchored oligo dT primer as 3:1 ratio in a 20 µL total reaction volume.

3.2.3 Real-time qRT-PCR (quantitative reverse transcription PCR)

The primers for germline-specific genes, *piwil1, piwil2, nanos2* and *nanos3* and reference genes, *ß-actin* and *elf1α,* were the same as in Chapter 2. The sequence of a germline-specific gene, *gfra1* and bcl2 family genes in Nile tilapia, *bcl-xLa, bcl-xLb, baxa, baxl, bcl-2* and *mcl-1a* were found in NCBI and primers were designed using Primer-BLAST in NCBI (Table 3.1). PCR products were cloned using pGEM T-easy vector systems (Promega, Madison, USA) and sequenced (GATC Biotech, Konstanz, Germany) to generate standard curves of each target gene to allow absolute quantification. Efficiency and R^2 of the standard curve were higher than 90% and 0.995, respectively, for all genes, and the melting curve of each gene was checked in each assay to assure the production of a single product (Table 3.1). Each reaction consisted of a total volume of 5 µL containing 1.3 µL of cDNA (1/5 diluted, approx. 5.2 ng RNA), 2.5 µL of SYBR green mix (Luminaris Color HiGreen qPCR master Mix, Thermo Fisher, Waltham, USA), 0.3 µM of each forward and reverse primer and MilliQ water up to 5 µL. The qRT-PCR reaction followed the protocol: 95°C for 10 min, followed by 45 cycles of 95°C for 15 sec and 60 – 62°C for 60 sec, and melting curve analysis using LightCycler 480 (Roche, Basel, Switzerland) and 384-well plate. The absolute copy numbers of every target gene were calculated based on its standard curve, and the relative mRNA levels of all the samples were calculated by normalising to the levels of a geometric mean of *ß-actin* and *elf1α* for each time point.

Table 3.1. Primers used for qRT-PCR and values of the standard curve

Oligo name		Oligo sequence (5'-3')	Size	Product length	Tm (°C)	Accession No.	Standard curve		
							Efficiency	R^2	Range
piwil1	F	ATGATCGTGGGCATCGACTGCTA	23	97	62	XM_003445546.2	0.968	0.998	$10\text{-}10^7$
	R	ACCACCTGCTCATGCTTTGGTTG	23						
piwil2	F	TGCCATCAAGAAGCTGTGCTGTG	23	73	62	XM_003445662.2	0.909	1.000	$10\text{-}10^6$
	R	CTGGGAAATTGTGCGGACGTTGA	23						
nanos2	F	CGGGAAAGTTTTCTGCCCCATCC	23	140	62	XM_005448855.1	0.944	0.999	$10\text{-}10^7$
	R	AGAACTTGGCCCCTGTCTCCATC	23						
nanos3	F	GGAGTGTGACATGAGCCGAGCTA	23	116	62	XM_005460553.1	0.903	0.996	$10\text{-}10^7$
	R	AACTCGTTAGTGCACATTCGCGG	23						
gfra1	F	AAGCGACCAAACAGCACGGTAAG	23	92	60	XM_003441935.3	0.924	0.999	$10\text{-}10^7$
	R	GTCTGTTTTCCACACAGCAGCCA	23						
bcl-xLa	F	GGCTTTATGACACCGGCACAACA	23	110	62	XM_003456961.4	0.911	0.995	$10\text{-}10^6$
	R	TGCACCTGGAGTATGAACAGGCA	23						
bcl-xLb	F	GGAGGATGGGACCGCTTTACAGA	23	109	62	XM_003442737.4	0.920	0.998	$10^4\text{-}10^7$
	R	GGAGGATGGGACCGCTTTACAGA	23						
baxa	F	CGGTCGGGGTTTTCTTGGCA	20	128	60	XM_019357746.2	0.981	0.999	$10\text{-}10^7$
	R	TCCTCCTGTGCCCATTTCCC	20						
baxl	F	GCTCGAGTCAGTTCAGACTGGGT	23	83	62	XM_003456558.5	0.950	0.999	$10\text{-}10^7$
	R	AAACTCCCTCGGGTCTCTGTGTG	23						
mcl-1a	F	TATGTCATTTGTAGCGAAGAGCCT	24	124	60	XM_013264682.1	0.906	0.998	$10\text{-}10^7$
	R	GCAGTTGTCCCTGCCCTTTT	20						
bcl-2	F	ACCCCTGCTCGTTAGGGTGA	20	140	60	XM_003437902.4	0.905	0.996	$10\text{-}10^6$
	R	CACTGCGCTGGTTGTAGCGA	20						
β-actin	F	ATCTACGAGGGTTATGCCCTGCC	23	118	62	KJ126772.1	0.934	0.996	$10\text{-}10^7$
	R	CTGTGGTGGTGAAGGAGTAGCCA	23						
elf1α	F	CTGGACAAACTGAAGGCTGAGCG	23	116	62	NM_001279647.1	0.967	0.999	$10^2\text{-}10^7$
	R	AAGTCTCTGTGTCCAGGGGCATC	23						

3.2.3 Testis histology

Testis samples were fixed in Bouin's solution, then dehydrated, cleared and impregnated with paraffin wax using an automated Tissue Processor. The samples were embedded in paraffin wax using a histoembedder (Leica UK Ltd., Milton Keynes, UK). The wax blocks were trimmed and sectioned using a Rotary microtome (Leica UK Ltd., Milton Keynes, UK) at 5 µm thickness. The sections were transferred onto slides, dried and left overnight in oven. On the following day, these slides were stained with haematoxylin

and eosin. Sagittal sections of testes were observed for 0 DD (n = 3 fish) and 1200 to 3800 DD samples (n = 6 fish/treatment/time point). Both sagittal and transverse sections were observed for 5200 DD testes samples (n = 12 fish/treatment).

3.2.4 Steroid hormone analysis

Blood samples (n = 6 fish/treatment/time point) were collected from the fish caudal veins and stored at 4°C overnight. Serum was collected after centrifugation (1400 xG for 10 min) and stored at −20°C as aliquots until use. 50 µl of each serum sample was mixed with 1 mL of ethyl acetate and centrifuged at 1500 rpm for 10 min at 4°C, and the supernatants were evaporated using a vacuum centrifuge at 30°C for 40 min, then the extract was dissolved in an assay buffer provided by the enzyme immunoassay kits (EIA) for 11-KT and T (Cayman, Ann Arbor, USA) (Chen *et al*., 2017), while serum samples were directly used to measure cortisol level by enzyme-linked immunosorbent assay (ELISA) (IBL Int., Switzerland) (Chabbi and Ganesh, 2017), according to the manufacturer's instructions. Intra- and inter-assay coefficients of variation (CVs) in 11-KT and T assays were less than 12.5% and 12.6%, respectively. The inter-CV in cortisol assay was 2.8%. Each assay was run with the serial dilution of unknown tilapia serum sample to confirm linearity of measurement by the assay (R^2 = 0.998 – 1) along with the standard curve samples to allow direct interpretation of hormone level in unknown samples (R^2 = 0.992 – 0.999).

3.2.5 Fertilisation rate and sperm quality analyses

At the end of the recovery period, between 5000 – 5200DD, fertilisation and sperm quality tests were conducted twice for each fish using a total of six independent batches of eggs (n = 2 egg batches/treatment). Semen from HT1, HT2 or CT males (n = 12 fish/treatment) were collected and stored on ice. Eggs of each batch collected from XX females were distributed into 12 Petri dishes, corresponding to one Petri dish for each semen sample. Semen collected from randomly picked CT (n = 2) were used to confirm egg viability (> 30% fertilisation rate in CT) in the fertilisation test of HT1 and HT2 groups. Eggs were fertilised *in vitro* by adding 1 µL of semen from each male. 1 µL of semen was chosen due to the minimum semen volume of 2 µL which could be collected from fish in the HT2 group. After 3 min of incubation, fertilised eggs were gently washed with the aquarium water and incubated in a tumbling egg system in isolation at 27 ± 1°C

(total of 72 batches over the period). The fertilisation rate was determined at blastula stage (4 hours post fertilisation) under a stereomicroscope.

Following the setup of the fertilisation test and within 2 hours of semen collection, sperm motility, duration of motility and spermatozoa density were analysed from each collected semen sample (1 µL) using a hemocytometer. The motility was scored based on the percentage of motile sperm in a semen sample after dilution with a scale between zero and five: zero, no motility; one, 1 – 20%; two, 21 – 40%; three, 41 – 60%; four, 61 – 80% five, 81 – 100% motile (Fauvel *et al.*, 1999). Distilled water was used as the dilution/activation medium. The time from the sperm activation to the end of sperm motility was recorded. Finally, spermatozoa density was estimated using immotile spermatozoa and a hemocytometer. The number of spermatozoa was counted in five of twenty five squares within the central counting area of two chambers, and it was carried out in triplicate for each sample. The sperm density was calculated by this following formula:

- Spermatozoa density (cell/mL) = average cells count in 5 squares from two chambers in triplicates $\times$ 5 $\times$ dilution factor $\times$ 10,000 (conversion factor from 0.1 mm^3 to mL).

3.2.6 Statistics

Data are presented as mean ± SEM. Statistical analysis was performed using Minitab 17 (Minitab Inc., State College, USA). Data was transformed when necessary to meet the homogeneity of variance and normal distribution. Significant differences in data between different developmental stages and temperatures (morphometric indices, expression level of germline-specific genes and Bcl-2 family genes and serum levels of 11-KT and T) were tested by two-way ANOVA, followed by Tukey's HSD test. Two-Sample t-test was used to evaluate the significant difference in sperm quality (motility, duration, density and fertilisation rate) ($p < 0.05$).

3.3 Results

3.3.1 Mortality and morphometric indices

During the first window of the heat treatment (up to 1200 DD), mortalities in CT, HT1 and HT2 were 2.5, 3.3 and 4.2%, respectively. There was no mortality in CT and HT1

between 1200 and 2100 DD, but 12.3% in HT2. During the last phase of the heat treatment (2100 to 3000DD), mortalities of 7.4 and 16.2% were observed in CT and HT2, respectively, and none in HT1.

Significant time, temperature and interaction effects were observed in body weight ($F = 147.83$, $p < 0.001$; $F = 103.62$, $p < 0.001$ and $F = 16.22$, $p < 0.001$, respectively) at 3000 and 5200 DD. At the end of the heat treatment (3000 DD), HT1 and HT2 mean body weight was 4.5 and 11.1 times lower than for CT, respectively (Table 2). HT1 and HT2 fish significantly gained weight during the recovery period; however, the weights remained significantly lower than that of control at 5200 DD (1.9 and 2.9 times lower, respectively).

For the GSI, significant time and temperature effects were observed ($F = 132.89$, $p < 0.001$ and $F = 10.89$, $p < 0.001$, respectively) without an interaction effect ($F = 1.96$, $p = 0.159$) at 3000 and 5200 DD. Overall GSI in HT1 ($0.23 \pm 0.09\%$) and HT2 ($0.17 \pm 0.08\%$) treatments was significantly lower than CT ($0.36 \pm 0.12\%$) (Table 2). In addition, overall GSI at 5200 DD ($0.43 \pm 0.09\%$) was significantly higher than at 3000 DD ($0.08 \pm 0.05\%$) (Table 2).

For the HSI, significant time, temperature and interaction effects were observed ($F = 8.00$, $p = 0.008$; $F = 3.66$, $p = 0.038$ and $F = 32.68$, $p < 0.001$, respectively) at 3000 and 5200 DD. HT1 and HT2 fish has significantly lower HSI than that of CT at 3000 DD (1.9 and 2.3 times lower, respectively), but HSI of HT fish became significantly higher than CT at 5200 DD (1.7 and 1.5 times higher, respectively for HT1 and HT2).

For the VSI, significant time, temperature and interaction effects were observed ($F = 5.55$, $p = 0.025$; $F = 35.45$, $p < 0.001$ and $F = 57.18$, $p < 0.001$, respectively) at 3000 and 5200 DD. VSI of HT1 and HT2 fish were 1.7 and 2.5 times lower than that of CT at 3000 DD, respectively, but became comparable to CT at 5200 DD.

Table 3.2. Morphometric indices (body weight, GSI, HSI and VSI) at 3000 and 5200 DD. Data are presented as mean ± SEM ($n = 6$ fish/treatment/time point). Superscripts denote statistically significant differences between treatments and sampling points ($p < 0.05$)

	3000 DD			5200 DD		
	CT	**HT1**	**HT2**	**CT**	**HT1**	**HT2**
Body weight (g)	226.7 ± 26.2^{ab}	50.1 ± 9.4^{d}	20.4 ± 2.5^{e}	350.7 ± 30.9^{a}	189.4 ± 21.6^{bc}	120.6 ± 8.9^{c}
GSI (%)	0.17 ± 0.09	0.05 ± 0.01	0.03 ± 0.00	0.56 ± 0.09	0.40 ± 0.07	0.33 ± 0.07
HSI (%)	2.96 ± 0.11^{a}	1.58 ± 0.16^{b}	1.26 ± 0.15^{b}	1.71 ± 0.15^{b}	2.83 ± 0.26^{a}	2.48 ± 0.19^{a}
VSI (%)	13.04 ± 0.65^{a}	7.54 ± 0.18^{b}	5.21 ± 0.15^{c}	7.01 ± 0.63^{bc}	8.68 ± 0.41^{b}	7.65 ± 0.24^{b}

3.3.2 Testis histology

Prior to the heat treatment at 0 DD (120 dah), the average weight, total body length and height were 3.1 ± 0.4 g, 5.8 ± 0.2 cm and 1.8 ± 0.1 cm ($n = 6$). All the testes weighed less than 0.01 g, but all three testes observed histologically had advanced stages of testicular germ cells and fish No. 2 & 3 contained well-structured seminiferous tubules and lumen filled with spermatozoa. (Fig. 3.2).
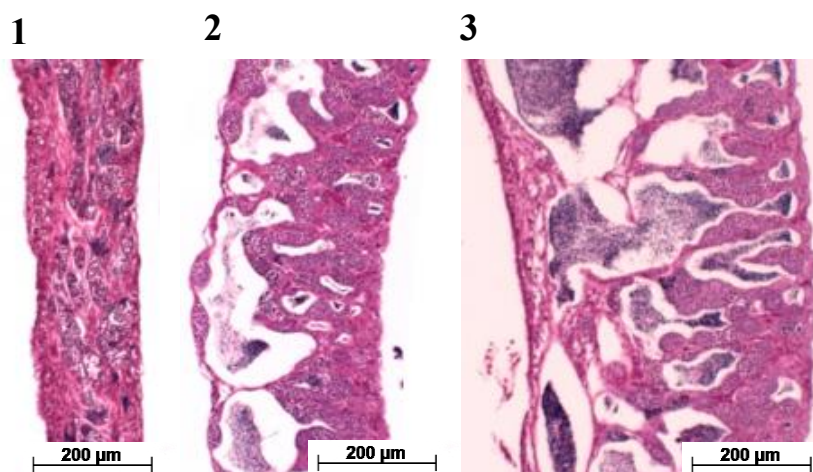
Figure 3.2. Histology of representative testes at 0 DD (120 dah) before the heat treatment ($n = 3$). All of three testes possess advanced stage of testicular germ cells and No. 2 & 3 contain well-structured seminiferous tubules and lumen filled with spermatozoa. All the testes weighed less than 0.01 g and fish No. 1, 2 and 3 were 1.9, 2.4 and 6.2 g in weight, 5.0, 5.5, 7.3 cm in body length and 1.4, 1.6 and 2.2 cm in body height, respectively.

Throughout the experiment, CT fish developed mature testes showing active spermatogenesis with well-structured seminiferous lobules (Fig. 3.3A-E). At 1200 DD, a reduction in the number of spermatids was observed in both HT1 and HT2 groups with increased vacuolation; however, spermatogonia and spermatocytes were still observed (Fig. 3.3F&K). Leydig cell hyperplasia was also observed with an apparent dense eosinophilic cytoplasm in HT treated fish from 1200 DD to 3800 DD (Fig. 3.3F-I&K-N). At 2100 DD, the level of atrophy and vacuolation increased in testes from both HT1 and HT2 fish, but the number of all stages of testicular germ cells was greatly reduced in HT2 than HT1 (Fig. 3.3G&L). By the end of the temperature treatments (3000 DD), testes of HT2 fish displayed increased severity of atrophy and vacuolation with testes from 7 out of 12 fish (58.3%) showing almost no testicular germ cells at any stages (Fig. 3.3M). In contrast, testis from all HT1 fish at 3000 DD presented all stages of testicular germ cells despite a severe loss of all stages of testicular germ cells and somatic cells (Fig. 3.3H). Testis from HT1 treated fish appeared to rapidly resume spermatogenesis with spermatozoa in the lumen of the seminiferous lobules by 5200 DD (Fig. 3.3I&J, Fig. 3.4B). However, two different groups in HT2 were observed at 3800 DD: one was lacking almost all stages of testicular germ cells (33%, 2/6) (Fig. 3.3N1), the other contained spermatogonia but no other spermatogenic stages (67%, 4/6) (Fig. 3.3N2). At the end of the recovery period (5200 DD), the majority of HT2 fish appeared to recover as evidenced by the presence of spermatozoa in the lumen (75%, 9/12) (Fig. 3.3O2), while a few (25%, 3/12) still appeared to lack spermatozoa at this time (Fig. 3.3O1, Fig. 3.4C No. 3, 5 & 12). Testes from HT2 males with less spermatozoa (Fig. 3.4C No. 1 & 11) or lack of spermatozoa (Fig. 3.4C No. 3, 5 & 12) were semi-transparent (Fig. 3.5C No. 1, 3, 5, 11 & 12) while the rest of the testes were whitish to pinkish in colour similar to CT and HT1 testes (Fig. 3.5). These HT2 males with less or lack of spermatozoa (Fig. 3.4C No. 1, 3, 5, 11 & 12) also had a significantly lower GSI than the rest of HT2 males at 5200 DD ($p < 0.05$) but there was no significant difference in their body weight (data not shown).
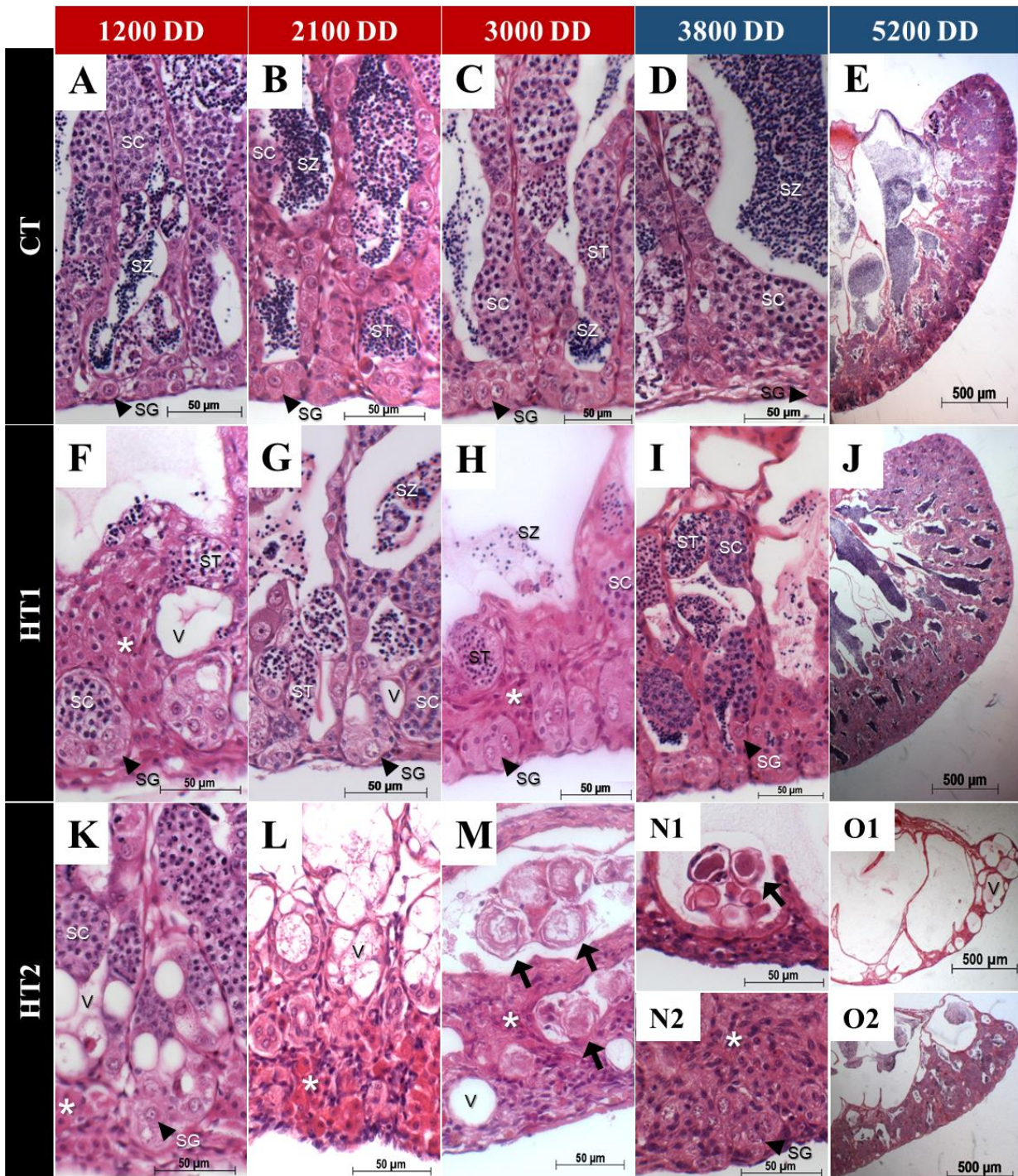
Figure 3.3. Histology of representative testes at 1200, 2100, 3000, 3800 and 5200 DD. (A-E) CT, fish exposed to 27°C, (F-J) HT1, fish exposed to 36°C (HT1, F-J) and (K-O) HT2, fish exposed to 37°C. N1 & O1 are representative histology of lack of testicular germ cells, while N2 & O2 represents re-proliferated teste in the recovery period. SG, spermatogonia (arrowheads); SC, spermatocytes; ST, spermatids; SZ, spermatozoa. Vacuolation (V), atrophy (arrow) and Leydig cell hyperplasia (*) are indicated.

A

Figure 3.4. Histology of transverse section of Nile tilapia testes at 5200 DD. (A) CT (*n* = 12), (B) HT1 (*n* = 12) and (C) HT2 (*n* = 12). Scale bar = 500 µm. (continued on next page)

B
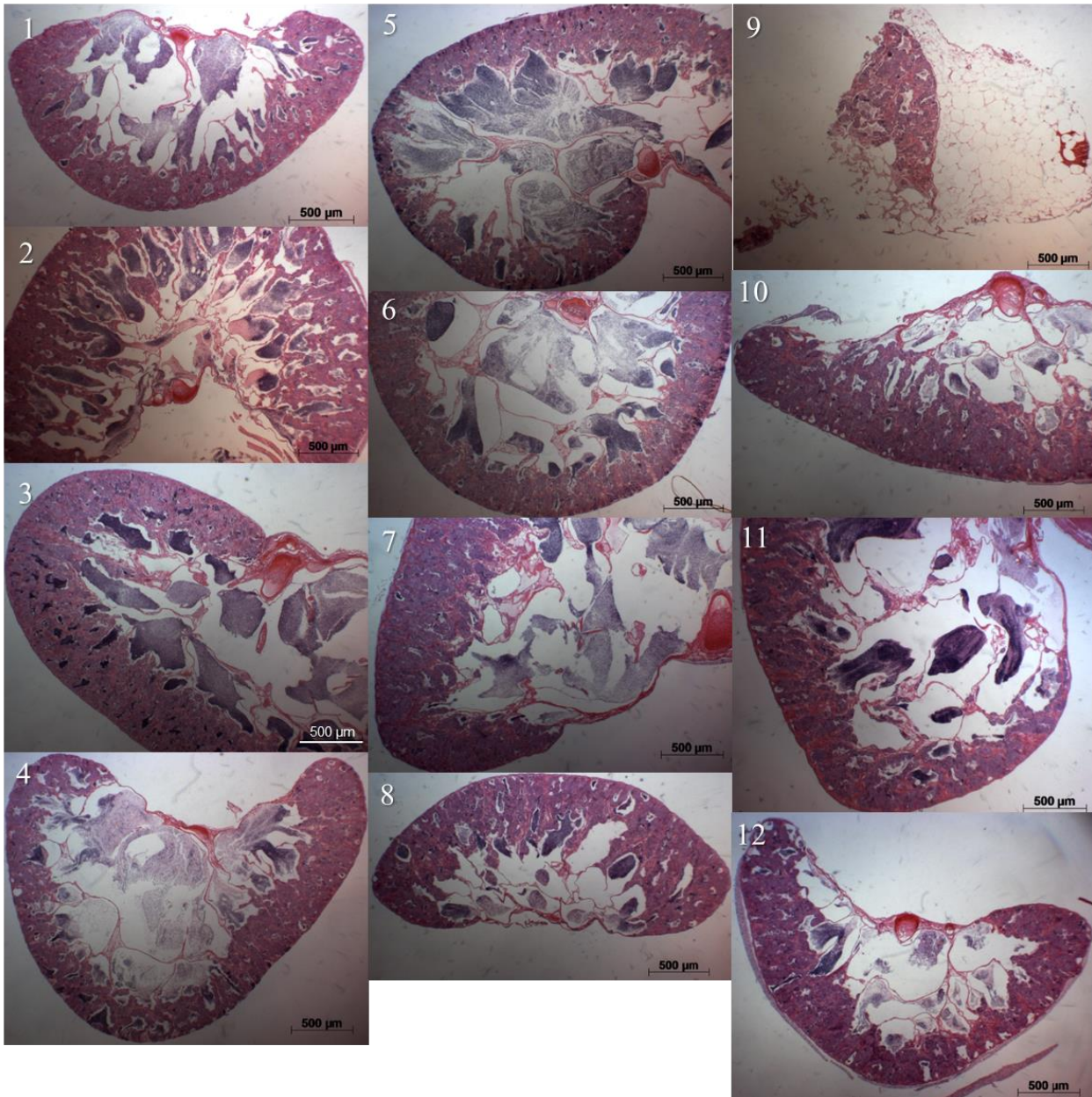
Figure 3.4. (continued from previous page)

C
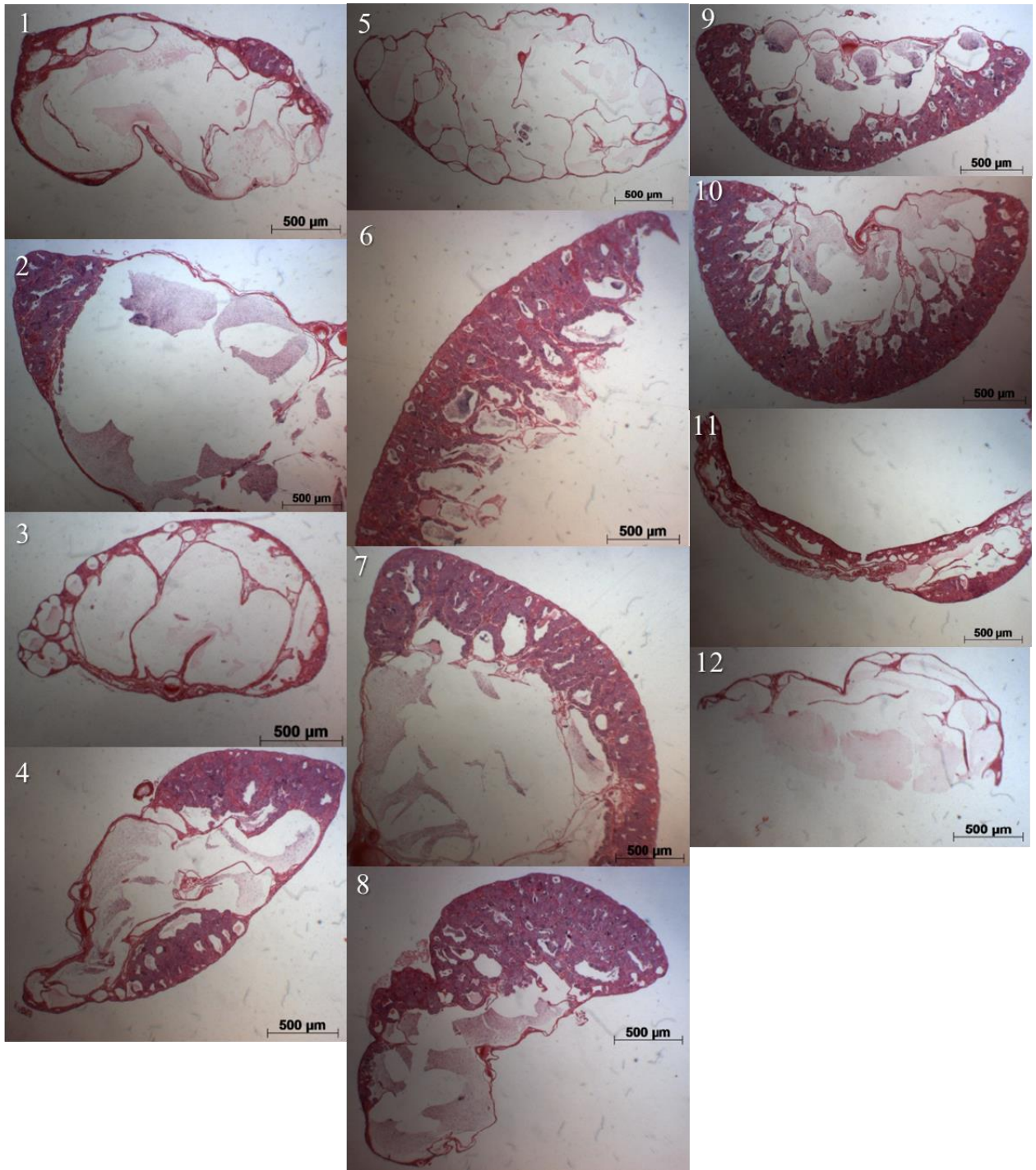
Figure 3.4. (continued from previous page)

A

Figure 3.5. Photographs of Nile tilapia testes at 5200DD. (A) CT (*n* = 12), (B) HT1 (*n* = 12) and (C) HT2 (*n* = 12). Scale bar = 10 mm. (continued on next page)

B

Figure 3.5. (continued from previous page)

C
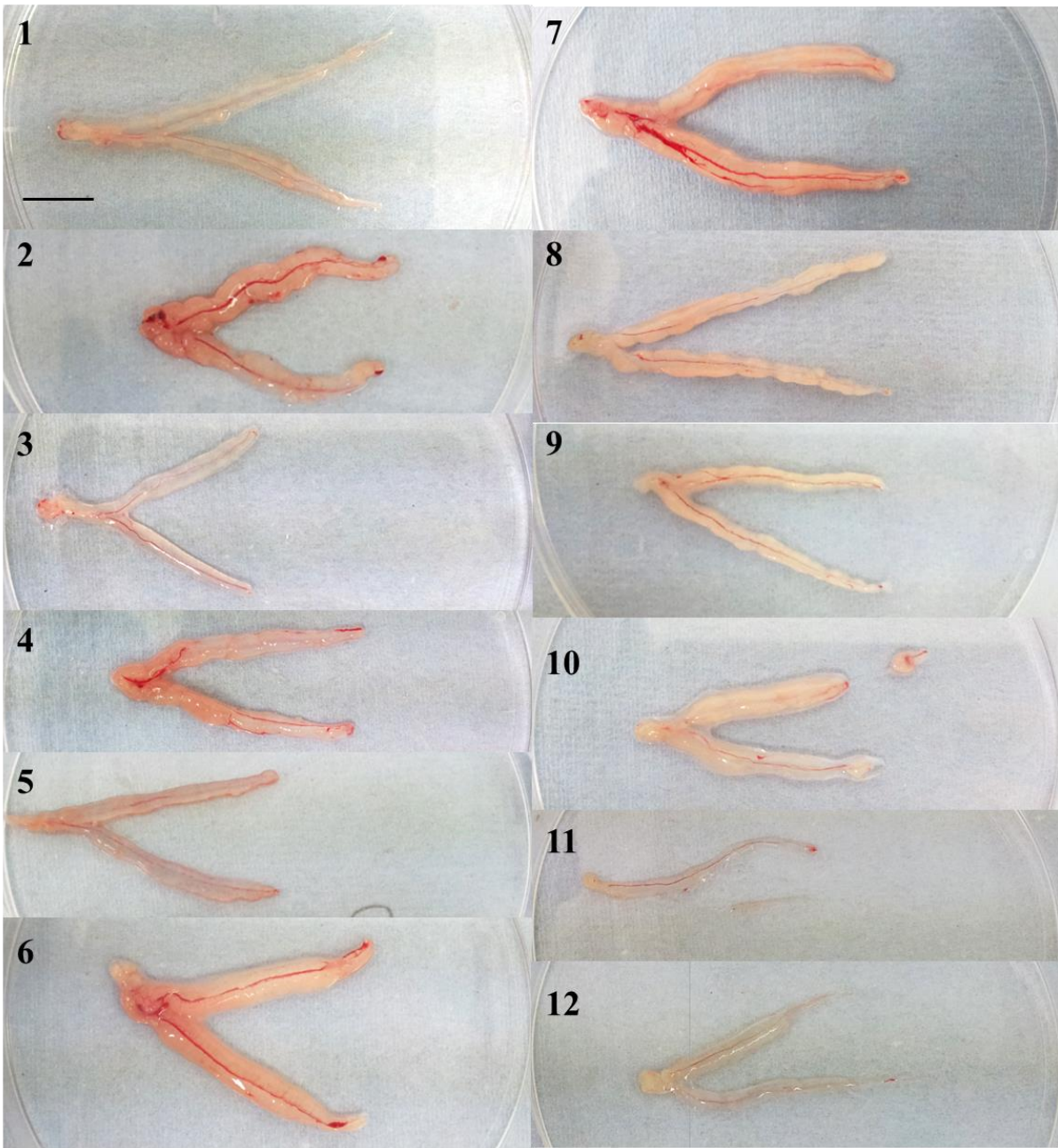
Figure 3.5. (continued from previous page)

3.3.3 Sperm quality

No significant differences in motility and duration were observed between treatments (Table 3.3). On the other hand, sperm densities of HT1 ($p = 0.037$) and HT2 ($p = 0.001$) were significantly lower than CT, showing 2- and 27-fold decreases, respectively. The fertilisation rate was significantly reduced in HT2, compared to CT ($p < 0.001$) and HT1 ($p < 0.01$) treatments, being 72.5% reduced compared to the CT treatment (Table 3.3).

Table 3.3. Sperm quality of milt collected from CT, HT1 and HT2 fish at the end of the recovery period (5200 DD). Data are presented as mean $\pm$ SEM ($n = 12$). Superscripts denote statistically significant differences between temperature treatments ($p < 0.05$)

Treatments	CT	HT1	HT2
Motility score (0-5)	4.8 ± 0.1	4.7 ± 0.1	4.3 ± 0.3
Motility duration (min)	2.6 ± 0.1	2.4 ± 0.0	2.5 ± 0.2
Spermatozoa density ($\times 10^9$) (cell/mL)	2.81 ± 0.55 [a]	1.47 ± 0.13[b]	0.10 ± 0.02[c]
Fertilisation rate (%)	41.5 ± 4.6[a]	35.0 ± 3.4[a]	11.4 ± 1.8[b]

3.3.4 Cortisol and sex steroid hormones

Serum cortisol levels were not significantly different between treatments during the whole experimental period with levels ranging from 124 and 275 ng/mL (data not shown). Significant time, temperature and interaction effects were observed in serum 11-KT ($F = 5.48$, $p = 0.001$; $F = 5.37$, $p = 0.007$ and $F = 8.87$, $p < 0.001$, respectively) while T levels showed significant interaction effects ($F = 4.90$, $p < 0.001$) without significant time and temperature effect. Serum 11-KT and T levels in fish from HT1 and HT2 were higher than that of CT at 1200 and 2100 DD with significant differences between HT1 and CT (Fig. 3.6). Levels then decreased to comparable levels than in CT fish at 3000 DD. During the first phase of the recovery window at 3800 DD, both serum 11-KT and T levels in HT1 and HT2 fish were significantly lower than that of CT fish, but levels returned to CT levels by 5200 DD (Fig. 3.6).
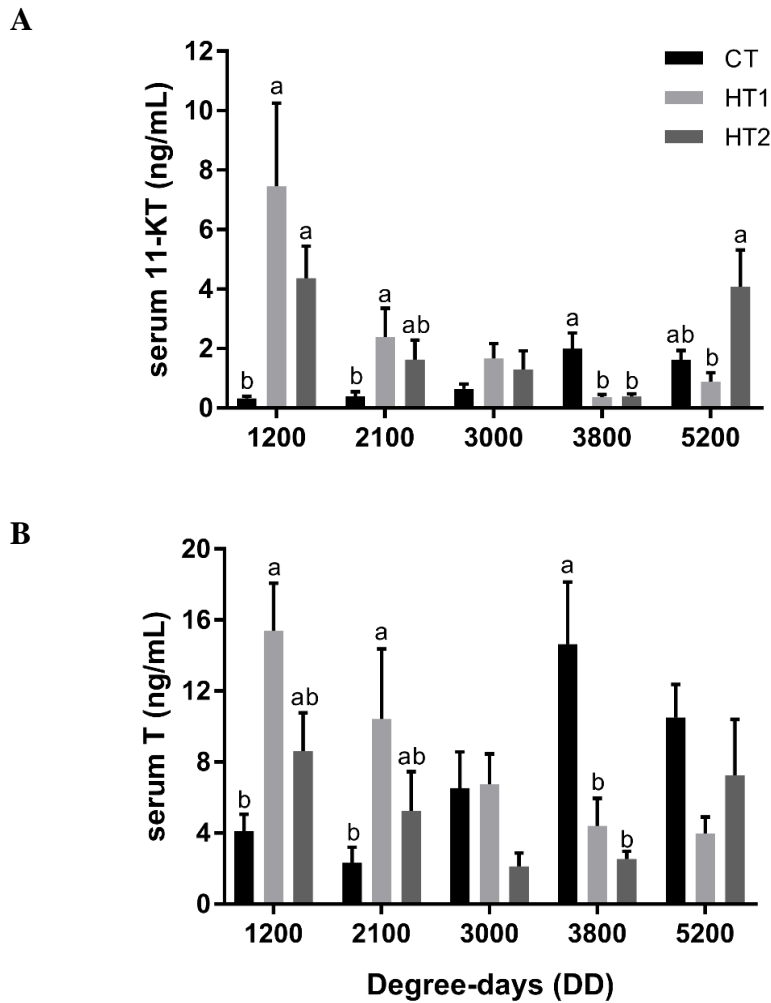
Figure 3.6. Serum sex steroid levels in Nile tilapia exposed to temperature treatments (CT, 27°C; HT1, 36°C; HT2, 37°C). (A) 11-ketotestosterone, 11-KT and (B) testosterone, T, levels at 1200, 2100 and 3000 DD during the treatment exposure window and 3800 and 5200 DD during the recovery period ($n = 6$). Data are presented as mean ± SEM. Superscripts denote significant differences between temperature treatment groups at each time point ($p < 0.05$).

3.3.5 Gene expression pattern

3.3.5.1 Germline-specific genes (*piwil*, *nanos* and *gfra1*)

Significant time, temperature and interaction effects were observed in relative gene expression levels of *piwil1* ($F = 47.91$, $p < 0.001$; $F = 145.07$, $p < 0.001$ and $F = 8.29$, $p < 0.001$, respectively) and *piwil2* ($F = 30.49$, $p < 0.001$; $F = 86.43$, $p < 0.001$ and $F = 5.27$, $p < 0.001$, respectively). Expression levels of *piwil1* and *piwil2* in testis of HT1 and

HT2 fish remained significantly lower than CT during the high temperature exposure window (1200 to 3000 DD) and the first phase of recovery (3800 DD), with HT2 being significantly lower than HT1 from 2100 to 3800 DD (Fig. 3.7A&B). During the last phase of the recovery period (3800 to 5200 DD), HT1 fish showed comparable expression levels to CT of both *piwil1* and *piwil2*; however, the levels remained significantly lower in HT2 compared to CT and HT1 at 5200 DD.

Significant time, temperature and interaction effects were observed in relative gene expression levels of *nanos2* ($F = 14.60$, $p < 0.001$; $F = 69.63$, $p < 0.001$ and $F = 8.22$, $p < 0.001$, respectively) and *nanos3* ($F = 3.37$, $p = 0.008$; $F = 35.65$, $p < 0.001$ and $F = 3.41$, $p < 0.001$, respectively). Expression levels of *nanos2* in both HT1 and HT2 fish were significantly lower than CT during the high temperature exposure window (1200 to 3000 DD) with HT2 being significantly lower than HT1 (1200 to 3000 DD) (Fig. 3.7C). On the other hand, only HT2 fish showed significantly lower expression levels of *nanos3* compared to CT during the heat treatment period (1200 to 3000 DD), while HT1 showed comparable expression levels to CT (Fig. 3.7D). During the recovery period (3800 - 5200 DD), expression levels of *nanos2* in HT2 fish remained significantly lower than CT and HT1, while the levels of *nanos3* increased in both HT1 and HT2 and reached the similar level to CT.

Significant time, temperature and interaction effects were observed ($F = 3.88$, $p = 0.003$; $F = 18.60$, $p < 0.001$ and $F = 2.76$, $p = 0.005$, respectively) in relative expression level of *gfra1*. Only HT2 fish showed significantly lower *gfra1* levels than CT fish between 1200 and 2100 DD (Fig. 3.7E). Following the return of all fish to control temperature, *gfra1* levels in HT1 fish were significantly higher than CT at 3800DD, and no significant differences were observed between treatments at 5200 DD.

Within HT2 males at 5200 DD, it was noted that testes from HT2 males with less or lack of spermatozoa (Fig. 3.4C No. 1, 3, 5, 11 & 12) showed significantly lower expression levels of *piwil1*, *piwil2*, *nanos2* and *nanos3* than the rest of HT2 male testes (Fig. 3.4C No. 2, 4, 6 – 10) ($p < 0.05$) but there was no significant difference in *gfra1* (data not shown).
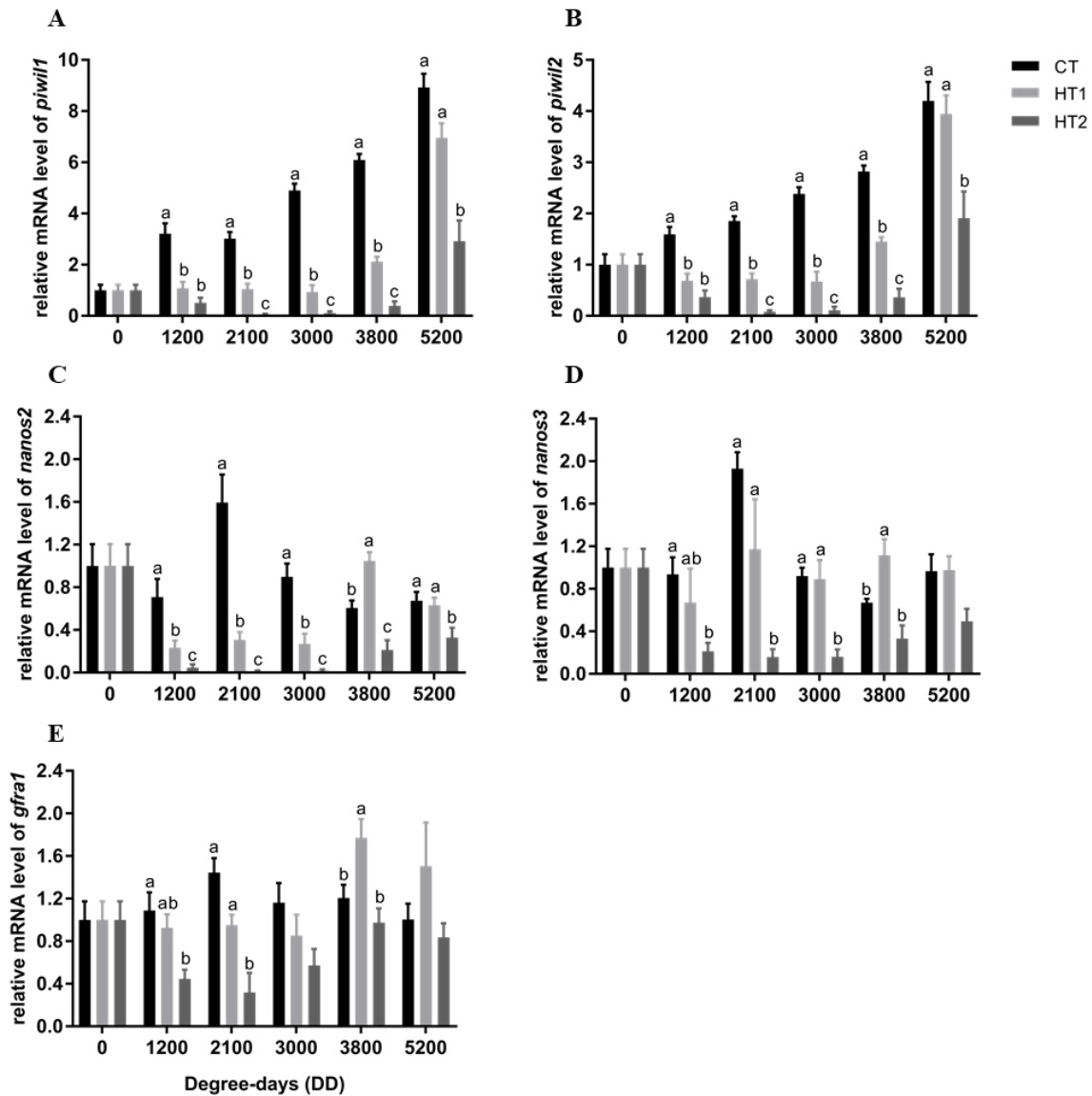
Figure 3.7. Normalised relative expression level of selected germline-specific genes in testis. (A) *piwil1,* (B) *piwil2,* (C) *nanos2,* (D) *nanos3* and (E) *gfra1*. Data are presented as mean ± SEM ($n = 6$). Superscripts denoted significant differences between treatments at a given time ($p < 0.05$). CT, 27°C; HT1, 36°C; HT2, 37°C.

3.3.5.2 Bcl-2 family genes

Significant time, temperature and interaction effects were observed in relative gene expression levels of *bcl-xLa* ($F = 7.96$, $p < 0.001$; $F = 22.59$, $p < 0.001$ and $F = 3.42$, $p = 0.001$, respectively) and *bcl-xLb* ($F = 37.82$, $p < 0.001$; $F = 172.97$, $p < 0.001$ and $F = 11.10$, $p < 0.001$, respectively). While the expression levels of *bcl-xLa* in HT2 fish were significantly lower than CT and HT1 only at 2100 and 3800 DD (Fig. 3.8A), the pattern of *bcl-xLb* mRNA level was similar to those of *piwil1* and *piwil2* (Fig. 3.8B, Fig. 3.7A & B). The expression levels of *bcl-xLb* were significantly lower in both HT1 and HT2 fish than CT fish from 1200 to 3800 DD with HT2 being significantly lower than HT1 from 2100 to 3800 DD, and levels remained significantly lower in HT2 compared to CT and HT1 at 5200 DD (Fig. 3.8B).

Significant time, temperature and interaction effects were observed in relative gene expression level of *baxa* ($F = 45.84$, $p < 0.001$; $F = 78.97$, $p < 0.001$ and $F = 6.73$, $p < 0.001$, respectively) and *baxl* ($F = 14.31$, $p < 0.001$; $F = 8.45$, $p < 0.001$ and $F = 2.46$, $p = 0.012$, respectively). The pattern of expression of *baxa* showed similarity to those of *piwil1*, *piwil2* and *bcl-xLb* (Fig.8B & C, Fig. 3.7A & B). Expression levels of *bcl-xLb* were significantly lower in both HT groups than CT from 1200 to 3800 DD, with HT2 being significantly lower than HT1 at 2100 and 3800 DD (Fig. 3.8C). At the end of the recovery period (5200 DD), expression levels of *baxa* in HT1 became comparable to CT fish, while levels in HT2 fish remained significantly lower than CT and HT1. On the other hand, no significant differences in expression levels of *baxl* between temperature treatment groups were observed except for *baxl* levels in HT1 at 1200 and 3800 DD which were significantly higher than CT (Fig. 3.8D).

Significant time, temperature and interaction effects were observed in relative gene expression levels of *bcl-2* ($F = 9.50$, $p < 0.001$; $F = 7.64$, $p = 0.001$ and $F = 3.28$, $p = 0.001$, respectively) and *mcl-1a* ($F = 5.40$, $p < 0.001$; $F = 4.47$, $p = 0.014$ and $F = 3.17$, $p = 0.002$, respectively). While no changes in the expression levels of *bcl-2* between temperature treatment groups were observed during the high temperature treatment, HT1 fish showed a significantly higher level of *bcl-2* than CT at 3800 DD and HT2 fish at 5200 DD than CT and HT1 during the recovery period. The expression levels of *mcl-1a* between temperature treatment groups were not different except for *mcl-1a* level in HT1 at 1200 DD which were significantly higher than CT and *mcl-1a* levels in HT2 at 3800 DD which were significantly lower than CT and HT1 (Fig. 3.8F).

Within HT2 males at 5200 DD, it was also noted that testes from HT2 males with less or lack of spermatozoa (Fig. 3.4C No. 1, 3, 5, 11 & 12) showed significantly lower expression levels of *bcl-xLb* and *baxa* and significantly higher expression of *bcl-2* than the rest of HT2 male testes (Fig. 3.4C No. 2, 4, 6 – 10) ($p < 0.05$) (data not shown).
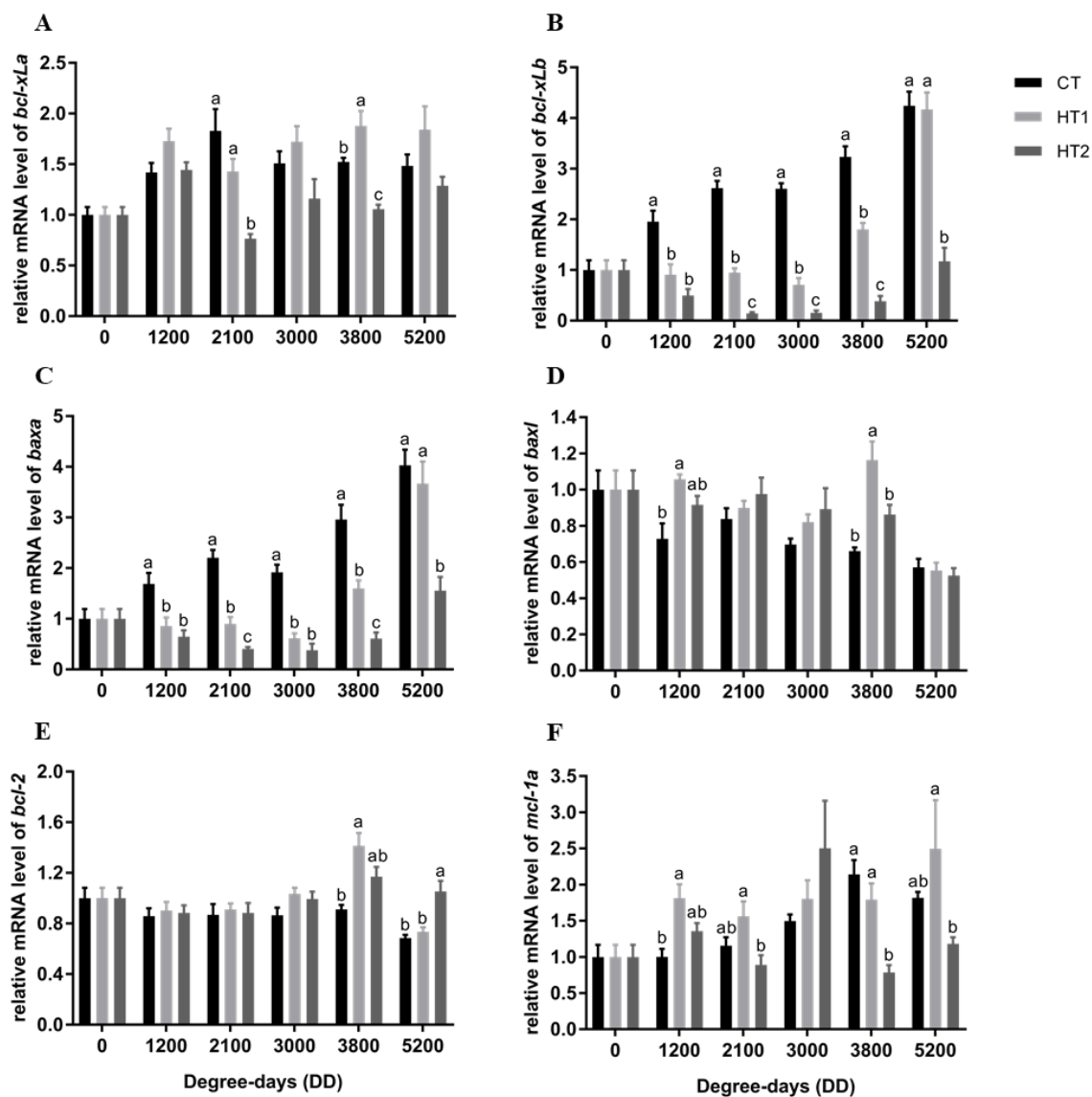
Figure 3.8. Normalised relative expression level of Bcl-2 family genes in testis. (A) *bcl-xLa*, (B) *bcl-xLb*, (C) *baxa*, (D) *baxl*, (E) *bcl-2* and (F) *mcl-1a*. Data are presented as mean ± SEM ($n = 6$). Superscripts denoted significant differences between treatments at a given time ($p < 0.05$). CT, 27°C; HT1, 36°C; HT2, 37°C.

3.4 Discussion

The aim of this study was to explore potential targets which play an important role in the survival or apoptosis of germ cells by utilizing high temperature as a means to induce germ cell death. Male tilapia exposed from 120 dah to high temperature treatments (36 and 37°C) for a duration of 3000 DD showed a significant loss of testicular germ cells compared to the control fish kept at 27°C. However, the HT treatments did not induce complete sterility in Nile tilapia. Importantly, the impact of HT on spermatogenesis studied at morphometric, histological, endocrine, gene expression and sperm quality levels, differed between 36 and 37°C with more acute and lasting effects being observed in fish exposed to 37°C.

HSI and VSI are often used as indicators of energy status, metabolism and condition in fish (Lambert and Dutil, 1997; Nunes *et al*., 2011; Wootton *et al*., 1978). The significantly reduced HSI and VSI at the end of the HT indicates the low energy status and metabolism in HT fish driven by the exposure to high temperature, accompanied by growth retardation compared to CT. HSI and VSI values in HT fish were reversible in the favorable environment showing a significant increase at the end of the recovery period. In addition, the significantly lower GSI values at the end of HT imply the suppression effect of HT on spermatogenesis, which was further confirmed by histological observation. Interestingly, GSI recovered at the end of the recovery regime showing high resilience in testicular development in this species. Such observed effects on retarded growth, lower HSI, VSI and GSI at the end of the HT window could be associated with reduced appetite and food intake as shown in fish under heat stress (Azaza *et al*., 2008; Brett, 1979; Handeland *et al*., 2008; Pandit and Nakamura, 2010); however, it must be acknowledged that feed consumption was not monitored in the current study. During the experimental design phase, there were concerns relating to the perceived severity of the HT2 treatment which was a driver of the reduced scale of the experiment. In association with the reduced growth and energy status, there was also an elevated mortality, primarily associated with the HT2 treatment. When such data are considered in combination, it can be concluded that heat treatment should not be considered to be a commercially viable strategy for sterilisation in fish.

Motility of sperm cells indicates the sperm fitness and viability (Stoss and Holtz, 1983) and a strong positive correlation between fertilisation rate and sperm motility is usually reported in teleost species (Lahnsteiner *et al*., 1998; Moccia and Munkittrick,

1987; Rurangwa *et al*., 2001). The duration of motility varies depending on species and water temperature, and in various teleosts species, it can last for up to 2 - 3 min (Billard, 1978; Burness *et al*., 2005; Chowdhury and Joy, 2001; Vladic and Jatrvi, 1997). In the current study, both motility and the duration of motility were not significantly different between HT and CT fish at the end of the recovery window, implying the fitness and viability of the sperm were similar to CT fish. However, both HT1 and HT2 male tilapia showed significantly reduced sperm density compared to CT with a more severe reduction in HT2. Moreover, HT2 showed a significantly reduced fertilisation rate compared to CT and HT1 fish. This link between sperm density and fertilisation rate has previously been reported in a number of teleost species (Aas *et al*., 1991; Tvedt *et al*., 2001). The reduced number of sperms in HT2 appeared to be through a loss of SSCs so that the number of spermatozoa they can produce was limited in HT male testes. However, HT2 did not result in sterility meaning that some SSCs were able to survive and proliferate to become functional spermatozoids. Whether HT2 treated individuals could return to a full competence given a longer recovery period is worthy of further research as it is equally important to understand germ cell recovery mechanisms to be able to truly define the efficacy of future alternative sterility inductions. Results from the current study contrast with previous findings in Mozambique tilapia in which a complete lack of testicular germ cells was reported in fish exposed to 37 ± 0.5°C for 50 days from 3 dah (Nakamura *et al*., 2015). These different effects could be explained by several factors including species and developmental stages. As the larval gonad has not fully developed the testicular somatic cells surrounding germ cells, larval germ cells are in a more vulnerable state due to the insufficient support and protection provided by testicular somatic cells. The essential role of Sertoli cells in germ cell survival was reported in mice where injured Sertoli cells caused apoptosis in testicular germ cells (Boekelheide *et al*., 2000). Even though HT failed to induce full sterility in male tilapia in this study, it gave an insight to study and understand the molecular mechanisms underlying germ cell loss by heat stress by enabling to analyse the morphometric indices, level of sex hormones in serum, testis histology and gene expressions through the heat treatment and recovery period, which were not possible in larval stage fish.

Germinal epithelium is the source of germline stem cells containing Sertoli cells and spermatogenic cells (Schulz *et al*., 2010). In Nile tilapia, temperature modulates Sertoli cell proliferation as shown in seasonal fish (de Alvarenga and de França, 2009). The number of Sertoli cells dictates the capacity of the testis to support testicular germ

cells as each Sertoli cell is able to support a fixed number of germ cells (Matta *et al*., 2002). Residual vacuolated Sertoli cells were seen in some atrophic germinal epithelium in HT fish, which probably accelerated apoptosis of testicular germ cells. There was also notable Leydig cells hyperplasia, which is commonly related to atrophic tubules in mammals (Greaves, 2012). In human, elevated androgen levels were reported in abnormal Leydig cells including hyperplasia (Wilson and Netzloff, 1983). In addition, LH overstimulation-induced Leydig cell hyperplasia caused a chronic increase of serum T levels in rats (Akingbemi *et al*., 2004). Therefore, the diffuse Leydig cells hyperplasia is likely to be associated with the significant increase of 11-KT and T serum levels in HT fish at 1200 and 2100 DD.

Stressors modulate fish reproduction through endocrine and paracrine pathways including androgens and cortisol by affecting the hypothalamus–pituitary–interrenal axis (Pickering *et al*., 1987; Goos and Consten, 2002; Ozaki *et al*., 2006; Pankhurst *et al*., 2008; Schreck, 2009). For example, temperature-induced cortisol modulated the transcript level of *11β-HSD2* responsible for the production of 11-KT in pejerrey at the sex differentiating stage (Lokman *et al*., 2002; Fernandino *et al*., 2012). In this study, HT had no apparent effects on plasma cortisol levels; however, it resulted in a significant increase of circulating plasma androgens (T and 11-KT) levels. The lack of HT effects on cortisol might be due to the chronic rather than acute nature of the stressor, although significant effects on growth may suggest otherwise. Another explanation could be the experimental design with only a few fish analysed and intra-specific variability which would have prevented the detection of differences. However, given the potential severity of the HT treatments, the number of fish under experimentation had to be maintained to a minimum. Even though there was no clear link between cortisol and sex steroid hormones in this study, the result implied the thermo-sensitivity of 11-KT. It is also demonstrated in Senegal sole (*Solea senegalensis*), where the change of thermocycle during sex determination and differentiation suppressed the 11-KT and T levels in muscle extract, causing a high proportion of females (Blanco-Vives *et al*., 2011). Higher 11-KT and T levels were also reported in heat-treated Mozambique male tilapia (Nakamura *et al*., 2015), suggesting a loss of germ cells may cause a loss of control in the production and release of androgens. Previous studies showed high plasma androgen levels can cause irreversible damage to SSCs and especially when germ cells are damaged by external factors (*e.g.* cytotoxic treatment, irradiation) they are more vulnerable to excessive androgen (Dohle *et al*., 2003; Meistrich *et al*., 2003). Excessive 11-KT related

impaired spermatogenesis was also reported in *figla* (*factor in the germ line, alpha*)-over-expressed XY tilapia (Qiu *et al*., 2015). Taken together, it is postulated that the overabundance of 11-KT and T induced by high temperature might accelerate the loss of testicular germ cells. Further investigation is required to understand the temperature-mediated control of androgen production and its effect on germ cell survival.

In contrast to previous work in tilapia (Nakamura *et al*., 2015; Pandit *et al*. 2015) this study examined the impact of temperature-induced germ cell loss at the molecular level. Both *piwil1* and *piwil2* have previously been reported to be expressed in various stages of testicular germ cells except for mature sperms in Nile tilapia (Xiao *et al*., 2013). In CT fish, both *piwil1* and *piwil2* expression levels gradually increased during the study in tune with the progression of the spermatogenesis in the CT fish. The significantly reduced levels of *piwil1* and *piwil2* in fish exposed to high temperature treatments are associated to the suppression of the spermatogenesis and the testicular germ cells reduction, which were more acute in fish exposed to HT2. The expression levels of both *piwil1* and *piwil2* in HT2 fish increased significantly during the recovery period compared to the levels during the heat treatment, but the overall reduced recovery in the HT2 treatment reflects the observed reduced functionality of the testis/sperm at the end of the study.

Unlike *piwil1* and *piwil2* which are expressed in various stages of testicular germ cells (Xiao *et al*., 2013), *nanos2*, *nanos3* and *gfra1* are known as SSC-specific markers in mammals and fish (Bellaiche *et al*., 2014; Suzuki *et al*., 2009). The patterns of *nanos2*, *nanos3* and *gfra1* in CT fish were not significantly changed during the study period, implying that there was no significant change in the number of SSC and/or spermatogonia in this period. Interestingly, the expression levels of *nanos3* and *gfra1* were reduced only in HT2 fish, while *nanos2* was downregulated in both HT1 and HT2 fish. Suzuki *et al*. (2009) reported differential expression of *gfra1*, *nanos2*, *nanos3* and *ngn3* (*neurogenin3*) within undifferentiated spermatogonial stages and proposed that SSCs are *gfra1* & *nanos2* positive and *nanos3* & *ngn3* negative, while *nanos3* & *ngn3* are widely expressed in undifferentiated spermatogonia in mice. Likewise, Lacerda *et al*. (2013) reported both *gfra1* and *nanos2* as SSC markers in Nile tilapia with *gfra1* being expressed exclusively in single type A undifferentiated spermatogonia (presumptive SSCs), while *nanos2* is also expressed in type A differentiated spermatogonia. Therefore, the decrease of *gfra1* expression levels in fish exposed to HT2 may reflect a significant reduction of single type A undifferentiated spermatogonia which only occurred in HT2

but not in HT1. *nanos3*-expressing spermatogonia were also significantly decreased in fish exposed to HT2 but not HT1. Yet, the role of *nanos3* in testis in this species has not been reported and further investigation is required to localize the expression of this transcript and determine its role in spermatogenesis. In addition, the expression levels of *nanos2* suggest that type A differentiated spermatogonia might have been significantly reduced by both temperature treatments, more so in fish exposed to HT2. These differences in the relative expression patterns of *nanos2*, *nanos3* and *gfra1* between fish exposed to HT1 and HT2 in this study support a more acute effect of HT2 than HT1 on the survival of different stages of spermatogonia in Nile tilapia.

Apoptosis is a process of programmed cell death which is essential for normal development and maintenance of normal cellular homeostasis in metazoan (Danial and Korsmeyer, 2004). Once apoptosis has begun, cell death is inevitable; therefore, the initiation of apoptosis is firmly regulated (Böhm and Schild, 2003). Apoptosis is initiated by two pathways, intrinsic and extrinsic pathways (Chauhan *et al*., 1997). The extrinsic pathway is mediated by the binding between extracellular death ligands and cell-surface death receptor, which activates caspase 8, and then triggers apoptosis of the cells (Wajant, 2002). In the intrinsic pathway, Bcl-2 family proteins interact with mitochondria causing the release of cytochrome c from mitochondria into the cytosol, which results in the formation of a complex between Apaf-1 and Caspase 9, leading to the activation of the caspase cascade (Danial and Korsmeyer, 2004). In the intrinsic pathway, a change in ratio between prosurvival and proapoptotic Bcl-2 family proteins is the decisive step for the initiation of apoptosis (Jia *et al*., 2007; Oltval *et al*., 1993; Rucker *et al*., 2000). This study profiled apoptosis-related genes to elucidate the apoptosis mechanism in germ cells under HT. *bcl-xL* are known to be prosurvival genes and their downregulation has been reported to cause the release of cytochrome c and activation of caspase cascades (Brockhaus and Brüne, 1999). In this study, long-term exposure to high temperature induced the downregulation of *bcl-xLb* but not *bcl-xLa*, which suggest differential functions of these genes under heat stress. Bax genes can be separated into three clades, *baxa*, *baxb* and *baxl* and *baxa* is clustered with mammalian homologs, while *baxb* and *baxl* are absent in amphibians and mammals (Li *et al*., 2017a). In Nile tilapia, *baxa* and *baxl* genes which were found in different chromosomes, LG4 and LG1 respectively, were investigated in the present study. The heat treatment appeared to suppress the expression of *baxa* but not *baxl*, suggesting potential functional differences between them under heat stress. Among six Bcl-2 family genes investigated in this study, *bcl-xLb* and *baxa*

expression levels were impacted by a long-term high temperature treatment in Nile tilapia, while *bcl-xLa, baxl, mcl-1a* and *bcl-2* were not significantly affected. Given that the imbalance between prosurvival and proapoptotic Bcl-2 family proteins is the key factor in apoptosis (Jia *et al*., 2007; Oltval *et al*., 1993; Rucker *et al*., 2000), it can be speculated that the significant decrease of *bcl-xLb* and *baxa* in HT1 and HT2 compared to CT during the heat treatment might have disrupted the rheostat of prosurvival/proapoptotic Bcl-2, which then in turn caused severe testicular germ cell loss through apoptosis. However, confirmation at the protein level is needed to confirm the heat stress-induced apoptosis of testicular germ cells in Nile tilapia.

Overall this work has provided a new perspective, at the molecular level, of mechanisms underlying testicular germ cell loss under heat stress. Importantly, this research has built on previous work of Chapter 2 and suggested the functional importance of *piwil* genes, in particular, in relation to germ cell loss and proliferation which prioritises them over the other candidates as a potential target for gene editing to induce sterilisation of Nile tilapia. At the same time this work looked at the molecular regulation of apoptosis in tilapia and has highlighted potential subfunctionalisation of gene roles within the Bcl-2 gene family that should be further investigated. High temperature treatment clearly impacted on spermatogenesis in Nile tilapia with the acuteness of the response being apparently temperature-dependent; however, given the observed physiological impacts on growth, energy status and survival, it can be concluded that thermal induction of sterility is not a commercially viable approach to be taken forward.

CHAPTER 4

OPTIMISATION OF CRISPR/CAS9 SYSTEM
IN NILE TILAPIA

4.1 Introduction

A remarkable improvement in genome editing technology in the recent year such as ZFN, TALEN and CRISPR/Cas9 system enables efficient and precise editing of the target DNA (Kim and Kim, 2014). The application of genome engineering techniques to understand the mechanisms that regulate the germ cell development opens new promising avenues to develop methods to control sexual maturation and mitigate associated detrimental effects in fish. To date, reproductively sterile fish induced by genome editing methods such as CRISPR/Cas9 system were reported in fish including *nanos2/nanos3* KO tilapia (Li *et al*., 2014), *dnd* KO Atlantic salmon (Wargelius *et al*., 2016) and *dnd*-KO medaka (Sawamura *et al*., 2017). In Nile tilapia, CRISPR/Cas9 system has been used for gene functional analysis such as *dmrt6* (Zhang *et al*., 2014b), *aldh1a2* (*aldehyde dehydrogenase 1 family member a2*) (Feng *et al*., 2015), *cyp26a1* (*cytochrome P450 family 26 subfamily a member 1*) (Feng *et al*., 2015), *amhy* (*Y chromosome-specific amh*) (Li *et al*., 2015), *amhrII* (*anti-Müllerian hormone receptor type II*) (Li *et al*., 2015), *gsdf* (Jiang *et al*., 2016) and *sf-1* (Xie *et al*., 2016), providing a powerful tool for gene functional analysis.

CRISPR/Cas9 system can be used to investigate the function of germline-specific genes which have the potential to induce sterility in fish. To investigate suitable target genes to induce sterility in fish by disrupting PGC formation or maintenance, GP components and germline-specific gene regulators including four *nanos*, two *piwil*, *dnd1*, *vasa* and three *pum* genes were investigated (Jin *et al*., 2019). To minimise pleiotropy and maximise the efficacy of PGCs diruption, desirable candidate genes for sterilisation should be exclusively expressed in gonads and have no apparent roles during embryonic development apart from PGC maintenance and survival. Thus, the ontogenic expression patterns and their tissue distribution were investigated and based on the result, *nanos3*, *piwil1*, *piwil2*, *dnd1* and *vasa* were listed as potential candidate genes. Among them, CRISPR/Cas9 mediated gene KO revealed the pivotal roles in germ cell survival of *nanos3* in tilapia (Li *et al*., 2014) and *dnd1* in medaka and salmon (Sawamura *et al*., 2017; Wargelius *et al*., 2016). Meanwhile, traditional mutagenic approaches were taken to produce *vasa*, *piwil1* and *piwil2*-null mutant zebrafish and results suggested their essential function in gametogenesis but not in the formation of PGCs (Hartung *et al*., 2014; Houwing *et al*., 2007, 2008). However, those homozygous null mutations were produced from fertile heterozygous parents, thus, there is a possibility that in this process

functionally potent mutation could be overlooked due to reduced fertility. In addition, a site-specific functional analysis is difficult to achieve in a randomly formed mutagenesis library. For these reasons, a highly efficient and target-specific genome editing technique such as CRISPR/Cas9 is desirable for the functional analysis of genes related to reproductive competence in injected animals. Vasa is a well-studied germline cell marker which belongs to ATP-dependent RNA helicase of the DEAD box family (Hay *et al*., 1988; Yoon *et al*., 1997) and in Nile tilapia genome, *vasa* has at least three loci and each homolog has a number of predicted transcript variants (Conte *et al*., 2017; Fujimura *et al*., 2011); therefore, complete KO of all *vasa* genes may be difficult to achieve in this species. On the other hand, Piwi members are components of the gene regulatory network of germline cells, together with Tudor domain protein and DEAD box helicase, which maintain germ cell fate and pluripotency of germline stem cells (Ewen-Campen *et al*., 2010). In Nile tilapia, there are also two *piwil* genes (*piwil1* and *piwil2*) (Xiao *et al*., 2013). In Chapter 2, it was shown that both *piwil1* and *piwil2* transcripts are maternally provided, no obvious zygotic expression during embryogenesis was found, and they are expressed exclusively in gonads in adult tissues (Jin *et al*., 2019). In this study, *piwil2* was selected as a target gene based on its suggested important role in PGC maintenance and gametogenesis; however, data on the function of this gene is lacking in other species. The functional analysis of *piwil2* was only reported in zebrafish among teleost species, using a homozygous null mutant line derived from the zebrafish ENU (N-ethyl-N-nitrosourea) mutagenesis library which has random point mutations (Houwing *et al*., 2008). However, this approach can not be easily replicated in other teleost species due to the high cost and long time to generate homozygous progeny. In addition, specific regions of interest within the target gene cannot be edited and the study of single mutations may mislead the interpretation of the function as the truncated protein may still play a role (Reva *et al*., 2011; Prykhozhij *et al*., 2017). In this context, *piwil2* KO using the target specific genome editing technique, CRISPR/Cas9, would effectively reveal the function of the gene in this species.

The application of a successful genome engineering method required first the optimisation of the delivery system for CRISPR/Cas9 molecules to ensure the method is efficient and reliable. There are three types of delivery techniques: (1) viral carriers such as adeno-associated viral vectors (AAVs) and non-integrating lentiviral vectors (NILVs), (2) non-viral carrier such as peptide-based carrier (CPPs), polymeric carriers and liposomes and (3) physical methods such as electroporation and microinjection (Ain *et*

al., 2015). The most widely used delivery method for ZFN, TALEN and CRISPR/Cas9 molecules are physical methods, which are simpler, target-specific and highly reproducible, using microinjection for embryos and electroporation for cell lines (Ain *et al*., 2015). Therefore, in this study, microinjection into tilapia zygote was chosen as a standard delivery method to conduct the functional study for *piwil2* gene using CRISPR/Cas9. One of the major bottlenecks in the microinjection process in tilapia eggs is that the chorion of tilapia egg is hard to penetrate. In most fish species, penetrating the outer surface coat or chorion can be difficult and there are several different techniques reported in the literature to overcome this problem (Jiang, 1993). One is to remove the chorion by manual dissection as done in zebrafish or by enzymatic dechorionation using trypsin or protease, as used in goldfish, carp, zebrafish, loach and northern pike (*Esox lucius*) (Cioffi *et al*., 1994). These methods are available in some species that have sufficient structural integrity without the mechanical support of the chorion (Stuart *et al*., 1988 cited in Cioffi *et al*., 1994). Another method is through piercing of the chorion using a thick and strong needle so that the microinjection pipette can be inserted through the opening. This method is suitable in species which produce large eggs with a chorion resistant to enzymatic digestion such as rainbow trout and salmon (McEvoy *et al*., 1988; Pinkert, 1994); however, it is a time-consuming procedure. A third is to insert the microinjection pipette through the micropyle which sits on the top of the embryonic cell; this has been used in Atlantic salmon (Fletcher *et al*., 1988) and tilapia (Brem *et al*., 1988; Fujimura and Kocher, 2011; Rahman and Maclean, 1992). However, this approach is also highly time-consuming as it requires perfect positioning of each zygote and it is limited by the short time window of micropyle opening. Therefore, it is critical to optimise the microinjection system of the CRISPR/Cas9 molecules for Nile tilapia zygote to improve the efficiency of the gene editing method without impacting on survival.

The next step is to optimise the formula of CRISPR/Cas9 molecules. Cas9 is guided by dual RNA structure consisting of crRNA (CRISPR RNA) which has a complementary sequence to the target and tracrRNA (trans-activating CRISPR RNA) which partially base pairs with the crRNA (Deltcheva *et al*., 2011). These crRNA and tracrRNA can be merged into a sgRNA (Jinek *et al*., 2012), making the process simpler. Therefore, the CRISPR/Cas9 system requires two components: universal Cas9 and target-specific sgRNA, and the design of target-specific sgRNA is a critical step for successful genome editing. The target sequence of CRISPR/Cas9 is usually 20 nt in length and should be located immediately next to the 5' of the PAM, which is 5'-NGG-3' for Cas9 derived

from *S. pyogenes* (Mojica *et al*., 2009). 12 nt in 3' of the target sequence, called the seed sequence, next to the PAM, serves as an essential recognition site for the binding of Cas9/sgRNA complex (Cong *et al*., 2013; Jinek *et al*., 2012). Therefore, it is crucial to design the target sequence which does not have potential off-target sites where the seed sequence and PAM are present in the genome apart from the target loci to minimise potential off-targets. It is, therefore, recommended to test multiple sgRNAs for each target gene to select the sgRNA with the highest efficiency for on-target mutation as the on- and off-target effects can vary between sgRNA (Doench *et al*., 2016).

In tilapia, the CRISPR/Cas9 technique has been utilised for genome editing studies (Li and Wang, 2017), but there is an apparent lack of data on the optimisation of the technique to ensure high efficiency and good survival. In a recently published study (Li *et al*., 2014), different concentrations of Cas9 mRNA have been tested and results showed that 500 ng/µL of Cas9 with 50 ng/µL of sgRNA resulted in lower survival rate, but higher mutation rate compared to 100 or 300 ng/µL of Cas9 with 50 ng/µL of sgRNA. However, this study only tested different concentrations of Cas9 mRNA (100, 300 and 500 ng/µL) with a constant concentration of sgRNA (50 ng/µL) and lacked replication. Further studies are required to determine the optimal ratio of Cas9 and sgRNA which results in reduced toxicity and increased mutagenesis efficiency.

The overarching aim of this study was to optimise the application of type II CRISPR/Cas9 system to edit the *piwil2* gene in Nile tilapia with the long-term goal being to validate a method which can reproducibly create *piwil2* mutant tilapia to explore the functional role of *piwil2* on development of PGCs in the species. Within this work three aspects of the CRISPR/Cas9 methodology were explored: (1) the optimisation of the microinjection delivery system using an inert dye (phenol red) to test the impact of the microinjection on embryo viability and realistically scale subsequent experimental designs; (2) to validate GFP linked reporter constructs to help visualise PGC presence that could be used to interpret the physiological impact of subsequent gene editing experiments and (3) test the effect of sgRNAs (100, 150 and 250 ng/µL) and Cas9 (500 ng/µL) mRNAs ratios to validate an optimal microinjection formulation which reliably produces *piwil2* mutant Nile tilapia.

4.2 Materials and Methods

4.2.1 Microinjection system for Nile tilapia embryo

4.2.1.1 Production of microinjection pipette and back-loading

Microinjection pipettes were prepared from glass capillary with filament (GD-1, Narishige, Tokyo, Japan), which outer and inner diameters were 1 and 0.6 mm, respectively. Each glass capillary was placed in the micropipette puller (Model 773, Campden, Loughborough, UK) and pulled using the program named Short shank microinjection pipette (Table 4.1). The tip of the pulled needle was broken using sharp forceps under a microscope. The diameter and shape of the micropipette tips were checked under a microscope and tested to find the optimal form. When the broken end was not sharp enough and/or too thin, it was either broken again or ground using a microgrinder (EG-44, Narishige, Tokyo, Japan).

Table 4.1. Puller Program. Program No. 2, short shank microinjection pipette (Instruction manual of Campden Model 773 Micropipette Puller, p 20)

Stage	Heat intensity (units)	Heating time (sec)	Pulling force (units)	Pulling displacement (mm)
1	20	0	20	3.0
2	0	1.0	0	0
3	15	0	50	24

4.2.1.2 Preparation of holding pipette

Capillary tubes were bent to around 60 - 80 degrees at 1/5 position in length by flame and the shorter end was fire-polished to make the pore inner diameter at around 0.5 - 0.7 mm. The longer end was connected to a plastic hose (around 40 cm in length), which end was connected to a Pasteur pipette tip. Then, it was fixed on the pipette holder to be connected to a manipulator (Narishige, Tokyo, Japan) on the opposite side of the injecting needle (Fig. 4.1A&B). By controlling the pipette bulb, the embryo can be gently held or released.
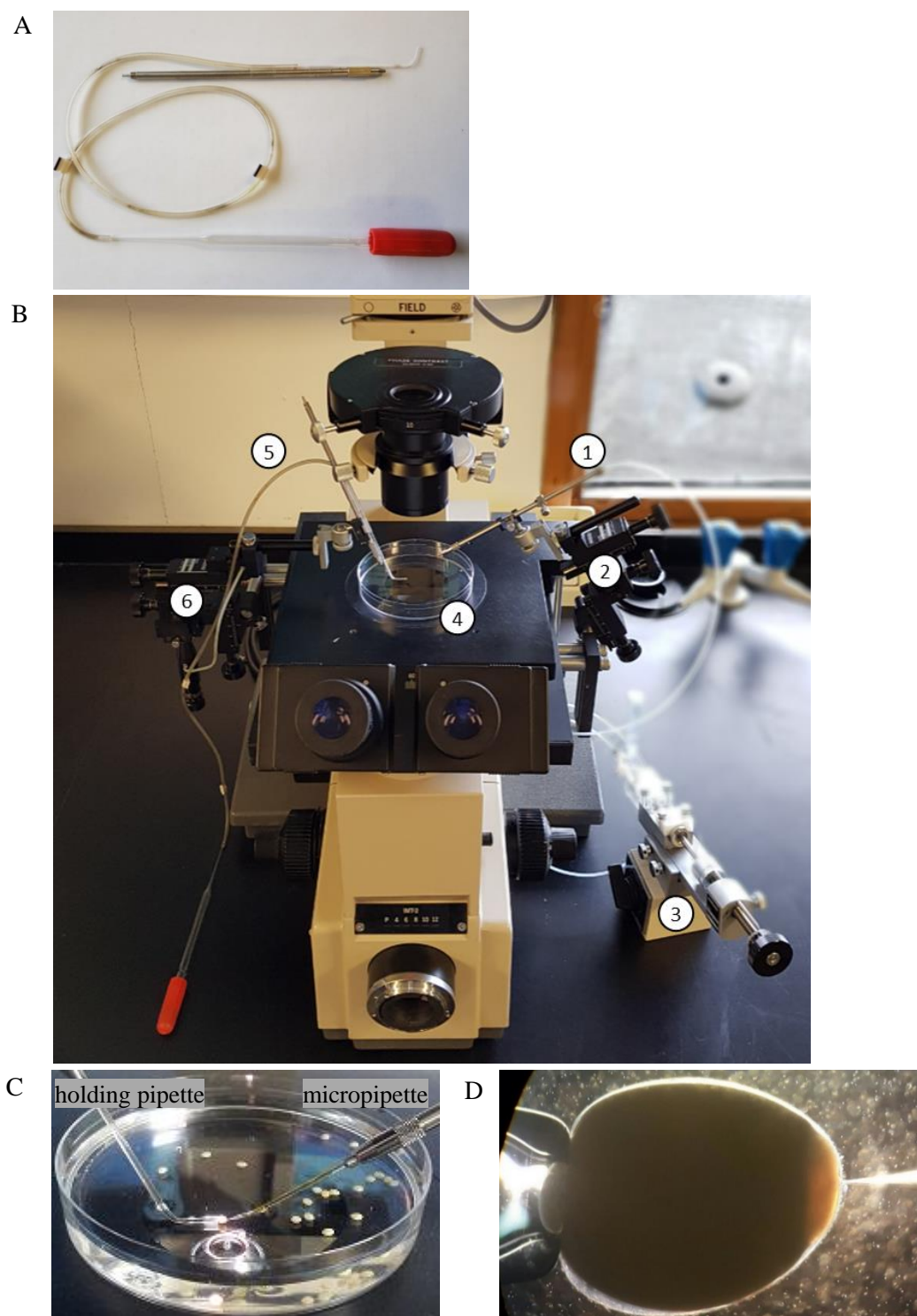
Figure 4.1. Microinjection system. (A) Holding pipette apparatus. (B) Assembled microinjection system. (C-D) An embryo is held by a holding pipette while a solution is injected into the embryonic cell under a microscope. 1, pipette holder; 2, manipulator for micropipette; 3, oil-based injector; 4, Petri dish; 5, holding pipette apparatus; 6, manipulator for holding pipette.

4.2.1.3 Handling of gametes and microinjection

To optimise the microinjection procedure, zygotes were produced from mature female (XX) and supermale (YY) tilapia (*O. niloticus*) held at the Institute of Aquaculture tropical aquarium. The eggs were washed with aquarium system water and kept at 24 – 26°C for up to 2.5 hours (hrs) before fertilisation. The milt was collected by glass capillary and stored immediately in a sterile test tube over ice for up to 2.5 hrs before fertilisation. Prior to fertilisation, eggs were gently washed with aquarium water and half of the egg batch (around 400 eggs) were fertilised in a Petri dish filled with aquarium water by adding 4 µL of milt. The fertilised eggs were washed off with aquarium water 3 minutes after the artificial fertilisation. Fertilised eggs were then kept at 20 - 22°C to extend the 1 cell stage for up to 2.5 hrs (Rahman and Maclean, 1992) for both control and injection groups. Thereafter, the remaining half of the egg batch was fertilised so that zygotes were available for microinjection for up to 5 hrs. Fertilised eggs at 1 cell stage were placed on a Petri dish filled with aquarium system water under an inverted microscope (IMT-2, Olympus, Tokyo, Japan). 3 – 5 µL of the injection solution was backloaded to the microinjection pipettes and filled with oil avoiding any air bubbles. Then, the pipette was connected to an IM-H1 pipette holder (Narishige, Tokyo, Japan) and oil-based IM-6 injector (Narishige, Tokyo, Japan). Each embryo was held by the holding pipette with its embryonic cell oriented towards the microinjection pipette. Then, the microinjection pipette penetrated both chorion and embryonic cells by using an M-152 manipulator (Narishige, Tokyo, Japan) and several nanoliters (approx. 4 nL) of the injection solution (phenol red or RNAs depending on experimental purpose) were injected into the embryonic cell while viewing through an inverted microscope. The injector used in this system manually controls the injection volume by turning the nob of the injector. Although it is not possible to strictly control the injection volume, it was aimed to inject approximately 4.2 nL of solution. The volume was determined by injecting into mineral oil placed on a micrometer (Rosen *et al*., 2009), where 4.2 nL has a diameter of 0.2 mm (Table 4.2). Following microinjection, the individual embryos were transferred to incubator using a dropper and a small spatula. As Cas9/sgRNA injected embryos were genetically modified, all experimental embryos/larvae were maintained in tanks designed for prevention of escapee (*e.g.* secured drain net) as approved by the Genetic Modification Safety Committee (GMSC) of the University of Stirling (Risk assessment US-GM042-IOA-AD). The control and injected eggs were incubated

according to standard practice for tilapia in round bottom recirculating incubators at 27 ± 1°C under 12L:12D photoperiod.

Table 4.2. Conversion table for calculating microinjection volumes ("Table 1. Conversion table for calculating microinjection volumes," 2010)

Diameter of drop (µm)	Radius of drop (µm)	Volume $(4/3[\pi r^3])$ (nL)
125	62.5	1.03
140	70	1.44
150	75	1.77
160	80	2.15
170	85	2.58
180	90	3.06
200	100	4.20
225	112.5	5.90
250	125	8.20
275	137.5	11.04
300	150	14.18
350	175	22.51

4.2.1.4 Phenol red microinjection

The viable scale of microinjection experiments was first tested using the inert dye phenol red. The aim was to test the impact of the physical microinjection of the eggs on mortality. Phenol red makes the visualisation of the microinjection location possible and is non-toxic during embryonic development (Layden *et al*., 2013). Fertilised eggs at 1 cell stage (approx. 30 embryos at a time) were placed in a Petri dish filled with aquarium system water under an inverted microscope (IMT-2, Olympus, Tokyo, Japan). Phenol Red (0.04%, visual aid) was back-loaded into the micropipettes using microloader tips (Eppendorf, Hamburg, Germany). Each embryo was held with a holding pipette and approx. 4 nL of 0.04% phenol red was injected into the embryonic cell. Eggs (including non-injected controls) were incubated as described above and survival rates were recorded at hatching (4 days post-fertilisation, dpf) and pre-first feeding stage (7 dpf, 3 dah) using three independent egg batches.

4.2.1.5 eGFP construct to label PGCs

GFP linked reporter construct that could visualise PGC presence would be a useful tool to interpret the physiological impact of subsequent gene editing experiments on PGCs (Gross-Thebing *et al*., 2017; Li *et al*., 2014). Therefore, four putative PGC labelling RNA constructs were designed to contain the ORF of eGFP (enhanced GFP) and 3'UTR of germline specific genes, *vasa*, *nanos3*, *piwil1* and *piwil2*. The ORF sequence of eGFP was obtained from pEGFP-C1 vector (Clontech laboratories, Mountain View, USA), which was a kind gift from Dr. Armin Sturm (Institute of Aquaculture, University of Stirling). First of all, the pEGFP-C1 vector was sub-cloned using JM109 comp cells (Promega, Madison, USA) and colonies were selected on LB agar with 30 µg/mL of kanamycin for further propagation in LB broth. The vector was used as a template to amplify ORF sequence of eGFP using Q5HiFi PCR master mix (NEB, Ipswich, USA) with a forward primer (5'-atggtgagcaagggcgaggagctgttcac-3') and a reverse primer which has restriction enzyme sites of *Afl*II, *Asc*I and *Nru*I at the end of the termination codon (5'-<u>tcgcgaggcgcgcccttaag</u>ttatctagatccggtggatc-3'). These three restriction enzyme sites were absent in ORF of eGFP as well as the 3'UTR of *vasa*, *nanos3*, *piwil1* and *piwil2* in addition to the pGEM T-easy vector sequence. Since Q5HiFi DNA polymerase with proofreading activity generates blunt-ended fragments, A-tailing of the purified PCR fragment was performed before the cloning into pGEM T-easy vector (Promega, Madison, USA). 3'UTRs of *vasa*, *nanos3*, *piwil1* and *piwil2* genes were amplified from Nile tilapia ovary cDNA using a forward primer designed after the termination codon and *Afl*II site was added to the 5' and a reverse primer designed close to the poly(A) tail and containing *Asc*I site at 5' (Table 4.3).

Table 4.3. Primer list used for production of PGC labelling RNA constructs

Gene	Accession No.	Forward primer (5'-3')	Reverse primer (5'-3')	Annealing temp (°C)
vasa	AB649032.1 (forward), AB649033.1 (reverse)	CTTAAGGAGCATCGCAGTCA CACAGCAATA	GGCGCGCCGAAACAGAACT ACAATTAGAT	58.9
nanos3	XM_005460553.3	CTTAAGTGAACCAGCAGGTG GCAAGG	GGCGCGCCCAAGAATAGCC ACAGAGAAA	59.3
piwil1	XM_013271587.1, XM_003445546.3	CTTAAGACCTTTAAGGATAA AAAGGCTTTGA	GGCGCGCCAGCACAACAAC AAGACT	53.5
piwil2	XM_003445662.3	CTTAAGTGAGTTATTCTTTG CCTTGATGCAG	GGCGCGCCTGAAAGCTGGA TTGGTGT	57.5

Each 3'UTR PCR fragment was purified using GeneJet PCR purification kit (Thermo Fisher, Waltham, USA) and cloned into pGEM T-easy vector. Both the plasmids of the 3'UTR's and eGFP were double digested with *Afl*II (NEB, Ipswich, USA) and *Asc*I (NEB, Ipswich, USA) and the double digested fragments were extracted from an agarose gel stained with ethidium bromide (EtBr) avoiding excessive exposure to UV, and ligated to each other using Quick ligation kit (NEB, Ipswich, USA). They were transformed into JM109 comp cells and cultured for plasmid extraction. The resultant sequence of each plasmid was confirmed by Sanger sequencing using universal primers, M13 forward and reverse (GATC Biotech, Konstanz, Germany). The plasmid was linearised by *Nru*I (NEB, Ipswich, USA) to be used as a template for *in vitro* transcription using HiScribe T7^{tm} high yield RNA synthesis kit (NEB, Ipswich, USA). The transcribed RNAs were purified using RNeasy mini kit (Qiagen, Hilden, Germany) and the size was checked on an agarose gel with ssRNA ladder (NEB, Ipswich, USA) (Fig. 4.2). The concentration was quantified using spectrophotometry (Nanodrop, Thermo Fisher, Waltham, USA) and diluted to 132, 176, 137 and 132 ng/µL for eGFP-3'UTR of *vasa, nanos3, piwil1* and *piwil2*, respectively (350 nM). The RNA constructs are shown in Fig. 4.2. Several nanoliters (approx. 4 nL) of 350 nM eGFP-3'UTR of *vasa*, *nanos3*, *piwil1* or *piwil2* RNAs was injected into the embryonic cell at 1 cell stage (three independent batches per eGFP-3'UTR construct). The injected live embryos/larvae ($n = 20 - 40$) were observed under a fluorescence microscope to detect GFP-expressing PGCs at a set interval of 24 hrs from 3 dpf to 5 dpf. Hatchlings were euthanised with Benzocaine to observe the fluorescence under the inverted microscope. Any constructs with higher than 80% labelling efficiency (% of embryos in a batch at a given life stage) would be considered as a successful PGCs labelling construct. All larvae then were killed by Schedule 1 method prior to being capable of first feeding.
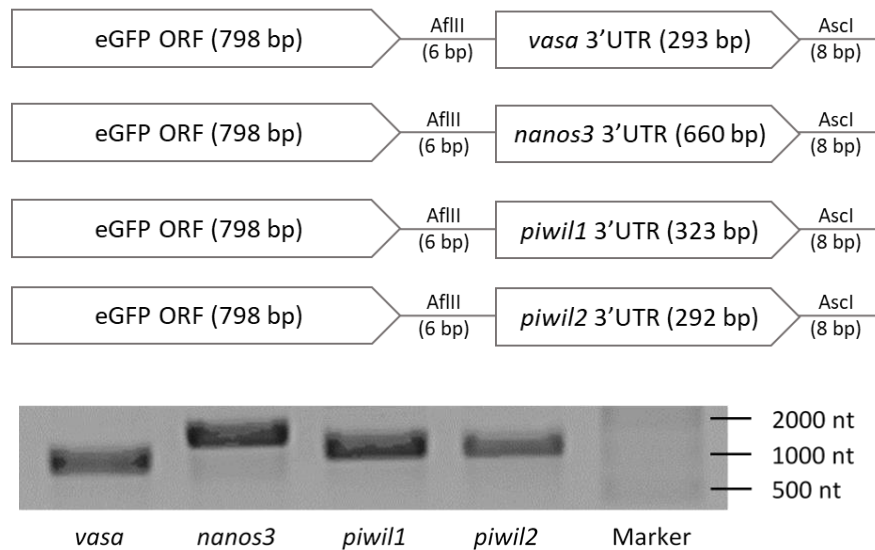
Figure 4.2. Schematic view of PGC labelling RNA constructs and the transcribed mRNAs. The size of transcribed mRNAs was 1169, 1536, 1199 and 1168 nt for eGFP-3'UTR of *vasa, nanos3, piwil1* and *piwil2*, respectively.

4.2.2 Design of sgRNA and preparation of sgRNA and Cas9 mRNAs

To knock out *piwil2* gene, two potential sgRNAs were designed using the online CRISPR RGEN tool (http://www.rgenome.net/) and Benchling (https://benchling.com) and the potential off-target sites of each candidate sgRNA was checked by searching similar sequences in tilapia genome using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), IoA BLAST server (http://139.153.176.55/blast/) and Cas-OFFinder (http://www.rgenome.net/cas-offinder/). Two sgRNAs were selected which had no apparent off-target alignment (no more than 12 nt were identical out of 20 nt sgRNA sequence together with PAM and were located in exon 24 (sgRNA1) and 21 (sgRNA2) on the conserved domain of PIWI (Fig. 4.3, Table 4.4). The secondary structure of sgRNAs was predicted using RNAfold (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold.cgi) to avoid sgRNA with a secondary structure which has potential constraints affecting the efficiency of sgRNA (Table 4.5).
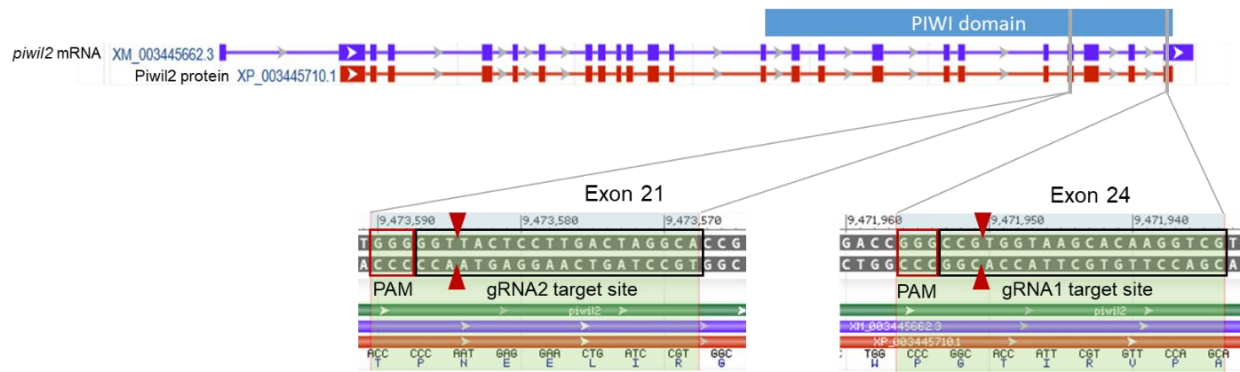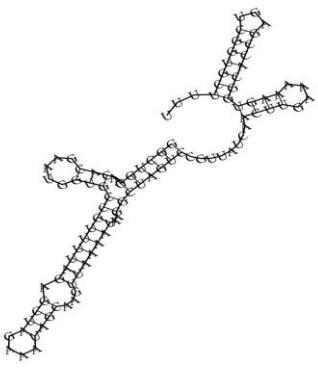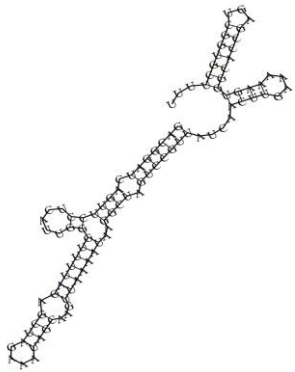
Figure 4.3. Target sites of Nile tilapia *piwil2* sgRNA1 and sgRNA2. Genomic structure of Nile tilapia *piwil2* gene (XM_003445662.3) and the protein (XP_003445710.1) are shown and sgRNA1 and 2 are located in 24th and 21st exons, respectively. Target site and PAM sequence are boxed in black and red, respectively. Arrowheads indicate the cutting site by Cas9 enzyme. *Piwil2* gene is located in LG12. The lengths of *piwil2* genomic sequence, mRNA and protein are 16.6 kb, 3.7 kb and 1067 aa.

Table 4.4. Potential off-target sites for *piwil2* sgRNAs used in this study

Bulge Type	Potential Off-target Sequence *piwil2* sgRNA2	Number of Mismatches		Chr	Accession No.	Position	Str	Gene
		Total	Bulge size					
DNA	crRNA: ACGG-ATCAGTTCCTCATTGGNGG DNA: ACGGCATCAGTTCgTCATTtGTGG	2	1	LG4	NC_031969.2	31418761	+	Intron of *noggin* (LOC100706105)*
RNA	crRNA: ACGGATCAGTTCCTCATTGGNGG DNA: ACGcATC-GTTCCTCATTaGAGG	2	1	LG3	NC_031967.2	32609461	+	
RNA	crRNA: ACGGATCAGTTCCTCATTGGNGG DNA: tCGGATCAGTTCCTgA-TGGAGG	2	1	LG18	NC_031982.2	882223	-	
RNA	crRNA: ACGGATCAGTTCCTCATTGGNGG DNA: ACtGAT-AGTgCCTCATTGGTGG	2	1	LG23	NC_031986.2	19247442	-	

Any sites consist of more than 12 nt identical to the seed sequence together with PAM were considered as potential off-target, but there were no such sites in Nile tilapia genome for either sgRNA1 or 2. In addition, 1 nt bulge (DNA/RNA) together with up to 2 nt mismatches in the target sequence was listed here. Dash indicates bulge sites, lowercase is the mismatched sequence and 12 nt of seed sequence is grey-shaded. The bulge type, total numbers of mismatches, size of the bulge, the target chromosome (Chr), the position of the cleavage site, the strand (str), the gene of the off-target site are shown. *, the off-target is located in the middle of an intron sized approx. 23 kb and the exons in both sides code 5'UTR of *noggin* (LOC100706105).

Table 4.5. Putative secondary structures of *piwil2* sgRNA1 and sgRNA2

	sgRNA1	sgRNA2
Graphical output		
Minimum free energy	-25.10 kcal/mol	-29.00 kcal/mol
Free energy of the thermodynamic ensemble	-26.70 kcal/mol	-30.02 kcal/mol
Frequency of the MFE structure in the ensemble[1]	7.43%	19.16%
Ensemble diversity[2]	18.61	5.86

[1]. The probability of a single structure in the Boltzmann weighted ensemble of all structures is given by $P = e^{\wedge}(-E/RT)/Q$. where E is the energy of the structure and Q is the partition function.

[2]. The ensemble diversity is the average base-pair distance between all structures in the thermodynamic ensemble.

4.2.3 Preparation of sgRNAs and Cas9 mRNA

  sgRNA template was produced by PCR (Fig. 4.4) (Shao *et al*., 2014). Briefly, the DNA template of the sgRNA scaffold was prepared by PCR using a primer pair for sgRNA scaffold (Table 4.6) and pT7-gRNA plasmid as template, which was a gift from Wenbiao Chen (Addgene plasmid #46759) (Appendix, Fig. S1A) (Jao *et al*., 2013). The product is universal for any sgRNA preparation, and it was used as a template to amplify DNA template for sgRNA through overlapping PCR using a primer pair introducing T7 promoter and target site on the 5' of sgRNA scaffold.

A

sgRNA scaffold reverse

pT7-gRNA T7 MCS sgRNA scaffold MCS M13F

sgRNA scaffold forward

PCR

B

sgRNA scaffold reverse

5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAAC-3'

sgRNA scaffold

5'-**GATCACTAATACGACTCACTATAGG**+NNNNNNNNNNNNNNNN+GTTTTAGAGCTAGAAAT–3'

T7 promoter + target site + sgRNA scaffold forward

PCR

C

T7 promoter target site sgRNA scaffold DNA template

IVT

D

target site sgRNA scaffold sgRNA

Figure 4.4. Schematic diagram to produce sgRNA using PCR approach to produce sgRNA template. (A) Preparation of the DNA template of the sgRNA scaffold through PCR using the pT7-gRNA plasmid as template. The product is universal for any sgRNA preparation. (B) Amplifying the DNA template for the sgRNA through overlapping PCR using the forward primer which has T7 promoter, target site and a 17-nt of 5' end of the sgRNA scaffold and the reverse primer designed on the 3' end of the sgRNA scaffold. (C) Purification of the DNA template for *in vitro* transcription of sgRNA. (D) Purification and quantification of the transcribed sgRNAs to be injected. T7, T7 promoter; MCS, multiple cloning site; M13F, M13 forward primer; grey shaded multiple 'N', target site sequence; IVT, *in vitro* transcription. This figure was modified from Shao *et al*. (2014).

Table 4.6. Primer list used for *piwil2* sgRNA production

Primer pair	Forward primer (5'-3')	Reverse primer (5'-3')	Annealing temp (°C)	Purpose
sgRNA scaffold	GTTTTAGAGCTAGAAAT AGCAAG	AAAAGCACCGACTCGGT G	53.7	Template of sgRNA scaffold
piwil2 T7 sgRNA1	GATCACTAATACGACTC ACTATAGGGCTGGAACA CGAATGGTGCCGTTTTA GAGCTAGAAAT	AAAAGCACCGACTCGGT GCCACTTTTTCAAGTTG ATAAC	71.6	Template of *piwil2* sgRNA1
piwil2 T7 sgRNA2	GATCACTAATACGACTC ACTATAGGACGGATCAG TTCCTCATTGGGTTTTA GAGCTAGAAAT	AAAAGCACCGACTCGGT GCCACTTTTTCAAGTTG ATAAC	70.9	Template of *piwil2* sgRNA2
M13 universal	GTAAAACGACGGCCAGT	AACAGCTATGACCATG	58	To amplify ORF of Cas9 from pT3TS-nCas9n

Grey shaded sequences are T7 promoter and underlined sequences are target of Nile tilapia *piwil2*.

The forward primer of *piwil2* T7 sgRNA1 or 2 contained three components; T7 promotor (5'-GATCACTAATACGACTCACTATAGG-3'), 20 nt of target sequence (orientation should be same with PAM) and 17 nt corresponding to the 5′ of the sgRNA scaffold (5'GTTTTAGAGCTAGAAAT-3'), while the reverse primer was designed on the 3' of the sgRNA scaffold (5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAAC-3') (Table 4.6). The PCR reaction of the sgRNA scaffold (80 bp), run in a total volume of 25 µL, consisted of 12.5 µL of MyTaq^tm 2 X Mix (Bioline, London, UK), 0.5 µL of 10 µM of each primer (sgRNA scaffold primer pair, Table 4.6), 0.5 µL of pT7-gRNA plasmid (157 ng) and Milli-Q water up to 25 µL. The PCR program was: 95°C for 1 min, followed by 25 cycles of 95°C for 15 sec, 53.7°C for 15 sec and 72°C for 30 sec, with a final extension at 72°C for 2 min, and conducted by Tgradient thermal cycler (Biometra GmbH, Göttingen, Germany). The PCR product was then purified using GeneJET PCR purification kit (Thermo Fisher, Waltham, USA) and the amplicon DNA at targeted size (80 bp) was confirmed by electrophoresis on a 1.2% agarose gel stained with EtBr and quantified using spectrophotometry (Nanodrop, Thermo Fisher, Waltham, USA). The purified sgRNA scaffold DNA (80 bp) was subsequently used as template DNA for the

amplification of the sg RNA templates (125 bp) which contain T7 promoter (25 bp), the target site (20 bp) and sgRNA scaffold (80 bp). The PCR reactions run in a total volume of 25 µL, consisting of 12.5 µL of MyTaq^{tm} 2 X Mix (Bioline, London, UK), 1 µL of 10 µM of each primer (*piwil2* T7 sgRNA1 or two primer pairs, Table 4.6), 1 µL of the purified sgRNA scaffold DNA (20 ng) and Milli-Q water up to 25 µL. PCR program was: 95˚C for 1 min, followed by 30 cycles of 95˚C for 15 sec, each annealing temperature (Table 4.6) for 15 sec and 72˚C for 30 sec, with a final extension at 72˚C for 2 min. The PCR product was purified using GeneJet PCR purification kit (Thermo Fisher, Waltham, USA) and the size (125 bp) was checked on the agarose gel. The purified PCR product was subsequently used as template for RNA synthesis using HiScribe T7^{TM} high yield RNA synthesis kit (NEB, Ipswich, USA). The transcribed RNAs were purified using RNeasy Mini Kit (Qiagen, Hilden, Germany) and quantified using spectrophotometry (Nanodrop, Thermo Fisher, Waltham, USA). The size (101 nt) and integrity of purified sgRNAs were checked by gel electrophoresis with Low range ssRNA ladder (NEB, Ipswich, USA). sgRNA solution was diluted and aliquoted for single use and stored at -20°C until use. The processed sequences of sgRNA1 and sgRNA2 is shown in Fig. 4.5B.

A

B

```
> sgRNA scaffold DNA (5'-3')
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT
CGGTGCTTTT


> piwil2 sgRNA1

GGCUGGAACACGAAUGGUGCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

> piwil2 sgRNA2

GACGGAUCAGUUCCUCAUUGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
```

Figure 4.5. sgRNA structure and sequences. (A) Structure of sgRNA and (B) the sequences (5'-3') of sgRNA scaffold DNA, *piwil2* sgRNA1 and sgRNA2. Underlined sequences are the targets of *piwil2* gene. The structure figure is from Prashant Mali (Church Lab) and Addgene (https://www.addgene.org).

Cas9 mRNAs were prepared using pT3TS-nCas9n, a gift from Wenbiao Chen (Addgene plasmid #46757) (Appendix, Fig. S1B) (Jao *et al*., 2013). The plasmids were sub-cloned and the sequences were verified by Sanger sequencing (GATC Biotech, Konstanz, Germany) (Fig. 4.6). This Cas9 sequence is type II CRISPR RNA-guided endonuclease Cas9 (*S. pyogenes*).

```
> pT3TS-nCas9n sequence (M13R – T3 promoter – Cas9 ORF – M13F)

AACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTAAGCTTGCTTGTTC
TTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGGCAGATCTAACTCGACGCCACCATGGCTTCTCCACCTAAGAAGAAGA
GAAAGGTGGGAAGCATGGATAAGAAGTATAGCATCGGCCTGGATATTGGAACTAACTCCGTGGGTTGGGCAGTGATTACAGA
CGACTACAAGGTCCCTAGCAAGAAATTTAAGGTGCTGGGTAACACCGACAGGCACAGCATCAAGAAAAATCTGATTGGAGCC
CTGCTGTTCGGTTCTGGAGAGACTGCCGAAGCAACACGCCTGAAAAGAACAGCAAGAAGGCGCTATACCAGAAGGAAGAATA
GAATCTGTTACCTGCAGGAGATTTTCTCTAACGAAATGGCTAAGGTGGACGATTCATTCTTTCATAGGCTGGAGGAAAGTTT
CCTGGTCGAGGAAGATAAGAAACACGAGCGCCATCCTATCTTTGGAAACATTGTGGACGAGGTCGCCTATCACGAAAAATAC
CCAACCATCTATCATCTGCGCAAGAAACTGGCTGACTCTACTGATAAAGCCGACCTGAGACTGATCTATCTGGCTCTGGCCC
ACATGATTAAGTTCAGGGGTCATTTTCTGATCGAGGGCGATCTGAACCCCGACAATTCCGATGTGGACAAGCTGTTCATCCA
GCTGGTCCAGATTTACAATCAGCTGTTTGAGGAAAAACCCTATTAATGCTTCCAGAGTGGACGCAAAAGCTATCCTGTCAGCC
AGGCTGTCCAAGTCACGCGAGACTGGAGAACCTGATTGCACAGCTGCCCGGAGAAAAGAGGAACGGTCTGTTTGGAAATCTGA
TCGCTCTGAGTCTGGGCCTGACTCCTAACTTCAAAAGCAATTTTGATCTGGCTGAGGACGCCAAACTGCAGCTGTCAAAGGA
CACATATGACGATGACCTGGATAACCTGCTGGCACAGATCGGAGATCAGTACGCTGACCTGTTCCTGGCTGCCAAAAATCTG
TCCGACGCAATCCTGCTGTCAGATATTCTGAGAGTGAACAGCGAGATTACAAAAGCTCCTCTGAGTGCCAGCATGATCAAGA
GATATGACGAGCACCATCAGGATCTGACCCTGCTGAAGGCTCTGGTCAGGCAGCAGCTGCCAGAGAAGTACAAGGAAATTTT
CTTTGATCAGTCCAAGAACGGCTACGCCGGTTATATCGACGGAGGCGCATCACAGGAGGAATTCTACAAGTTTATCAAACCT
ATTCTGGAGAAGATGGACGGAACTGAGGAACTGCTGGTGAAACTGAATAGAGAGGACCTGCTGAGGAAGCAGCGCACATTTG
ATAACGGTTCCATCCCACACCAGATTCATCTGGGAGAGCTGCACGCTATCCTGAGGCGCCAGGAAGACTTCTACCCCTTTCT
GAAAGATAACCGCGAGAAGATCGAAAAAATTCTGACCTTCAGAATCCCTTACTATGTGGGTCCACTGGCTCGCGGAAACAGC
AGATTTGCCTGGATGACTCGCAAATCCGAGGAAACCATTACTCCTTGGAACTTCGAGGAAGTGGTCGATAAGGGCGCCTCTG
CACAGTCCTTCATCGAGAGAATGACTAATTTTGACAAAAACCTGCCCAATGAGAAAGTGCTGCCTAAGCACTCCCTGCTGTA
CGAGTATTTCACTGTCTATAACGAACTGACAAAGGTGAAATACGTCACCGAGGGCATGAGAAAGCCAGCCTTCCTGTCAGGA
GAGCAGAAGAAAGCAATCGTGGATCTGCTGTTTAAAACCAACAGGAAAGTGACTGTCAAGCAGCTGAAGGAGGACTACTTCA
AGAAAATTGAATGCTTCGATTCCGTGGAGATCAGCGGAGTCGAAGACAGATTTAACGCAAGCCTGGGCGCTTACCACGATCT
GCTGAAGATCATTAAGGATAAAGACTTCCTGGACAACGAGGAAAATGAGGATATCCTGGAAGACATTGTGCTGACACTGACC
CTGTTTGAGGACAGAGGAATGATCGAGGAAAGACTGAAAACCTATGCTCATCTGTTCGATGACAAGGTGATGAAACAGCTGA
AGAGAAGGCGCTACACTGGCTGGGGTAGACTGAGCAGGAAGCTGATCAACGGCATTAGGGATAAACAGTCAGGAAAGACAAT
CCTGGACTTTCTGAAAAGTGATGGCTTCGCCAACCGCAATTTTATGCAGCTGATTCACGATGACAGTCTGACCTTCAAAGAG
GACATCCAGAAGGCTCAGGTGTCTGGACAGGGCCACTCCCTGCATGAGCAGATTGCAAACCTGGCTGGAAGCCCAGCCATCA
AGAAAGGCATTCTGCAGACAGTGAAATCGTCGATGAGCTGGTGAAAGTCATGGGCCATAAGCCCGAAAACATCGTGATTGA
GATGGCTCGCGAAAATCAGACAACCCAGAAGGGTCAGAAGAACAGTAGAGAGAGGATGAAAAGAATCGAGGAAGGCATTAAG
GAGCTGGGTAGCCAGATCCTGAAAGAGCACCCAGTGGAAAACACACAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGC
AGAATGGAAGAGATATGTACGTGGACCAGGAGCTGGATATTAACAGGCTGTCTGATTACGACGTGGATCATATCGTCCCCCA
GAGTTTCATCAAAGATGACAGCATTGACAACAAGGTGCTGACCAGGTCCGACAAAAACAGAGGAAAATCAGATAATGTCCCT
AGTGAGGAAGTGGTCAAGAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAAACTGATCACTCAGAGGAAGTTTGATA
ACCTGACAAAAGCAGAGCGCGGTGGACTGTCAGAACTGGACAAGGCTGGATTCATCAAGAGGCAGCTGGTGGAAACACGCCA
GATCACTAAACACGTCGCACAGATTCTGGATAGTCGCATGAACACAAAGTACGATGAGAATGACAAACTGATCAGAGAAGTG
AAGGTCATTACCCTGAAGAGTAAACTGGTCAGCGACTTTAGGAAAGATTTCCAGTTTTATAAGGTCCGCGAGATTAACAATT
ATCACCATGCCCATGACGCATACCTGAACGCCGTGGTCGGTACCGCACTGATCAAGAAATACCCAAAACTGGAGAGCGAATT
CGTGTACGGAGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGTCCGAGCAGGAAATTGGAAAAGCTACTGCCAAG
TATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACAGAGATCACCCTGGCCAATGGAGAAATCCGCAAAAGGCCCCTGA
TTGAGACAAACGGAGAGACAGGCGAAATCGTGTGGGACAAAGGCAGAGATTTTGCAACCGTGAGGAAGGTCCTGAGCATGCC
TCAAGTGAATATCGTCAAGAAAACTGAGGTGCAGACAGGCGGTTTCTCAAAAGAAAGTATTCTGCCAAAACGCAACTCTGAT
AAGCTGATCGCTAGAAAGAAAGACTGGGACCCTAAGAAGTATGGAGGCTTTGACTCTCCCACTGTGGCATACTCCGTCCTGG
TGGTCGCTAAGGTGGAGAAGGGCAAAAGCAAGAAACTGAAATCTGTCAAGGAGCTGCTGGGTATCACAATTATGGAGAGAAG
CTCTTTCGAGAAGAACCCAATCGATTTTCTGGAGGCCAAAGGTTATAAGGAAGTGAAGAAAGACCTGATCATTAAACTGCCC
AAGTACAGTCTGTTTGAGCTGGAAAACGGCAGGAAACGCATGCTGGCAAGCGCTGGAGAGCTGCAGAAAGGCAATGAACTGG
CCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCAAGCCACTACGAGAAGCTGAAAGGATCTCCAGAGGATAACGAACA
GAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCTAAAAGAGTGATC
CTGGCCGACGCAAATCTGGATAAGGTCCTGTCTGCTTACAACAAACACAGAGATAAGCCCATCAGGGAGCAGGCCGAAAATA
TCATTCATCTGTTCACTCTGACAAACCTGGGCGCACCTGCAGCTTTCAAGTACTTCGACACTACAATCGATAGAAAGAGGTA
CACCTCCACTAAGGAGGTGCTGGACGCTACACTGATCCATCAGAGTATTACCGGCCTGTACGAAACAAGGATTGACCTGTCT
CAGCTGGGTGGCGACTCACCGGTGAGATCTCCTAAGAAGAAGAGAAAGGTGTGACCGCGGATCTGGTTACCACTAAACCAGC
CTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAAAATGTAGCCA
TTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACATTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCCCCCCCC
CCCTGCAGGGCCAAGTCGGCCTCTAGAGGTACCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTA
C
```

Figure 4.6. Partial sequence of pT3TS-nCas9n (Addgene plasmid #46759) used for Cas9 mRNA production. Boxed, M13 Reverse primer; grey shaded, T3 promoter; underlined, ORF of Cas9 (4101 bp); grey shaded and boxed, M13 forward primer.

The Cas9 ORF template was amplified by PCR in a total volume of 25 µL, consisting of 12.5 µL of MyTaq™ 2 X Mix (Bioline, London, UK), 0.5 µL of 10 µM each primer (M13 universal primer pair, Table 4.6), 1 µL of pT3TS-nCas9n plasmid (70 ng) and Milli-Q water up to 25 µL. PCR program was: 95˚C for 1 min, followed by 25 cycles of 95˚C for 15 sec, 58˚C for 15 sec and 72˚C for 30 sec, with a final extension at 72˚C for 2 min. Then, the PCR product was purified using GeneJet PCR purification kit (Thermo Fisher, Waltham, USA) and the size (4593 bp) was checked on 1% agarose gel. The purified PCR product was subsequently used as a template for RNA synthesis using mMessage mMachine T3 kit (Thermo Fisher, Waltham, USA). The transcribed capped Cas9 mRNA were purified using RNeasy Mini Kit (Qiagen, Hilden, Germany) and quantified using spectrophotometry (Nanodrop, Thermo Fisher, Waltham, USA). The size (4542 nt) and integrity of purified Cas9 mRNA were checked by gel electrophoresis with ssRNA ladder (NEB, Ipswich, USA). The capped Cas9 mRNA solution was aliquoted for single use and stored at -20°C until use.

4.2.4 Injection of different ratios of Cas9 and sgRNA

Three different sgRNA concentrations (100, 150 and 250 ng/µL) combined with a single concentration of Cas9 (500 ng/µL) were tested for each sgRNA. The molar ratios of Cas9:sgRNA were 1:9.0, 1:13.4 and 1:22.4 for 500 ng/µL of Cas9 with 100, 150 and 250 ng/µL of sgRNA, respectively. The process was repeated for a given construct dose until data was collected from a minimum of three independent egg batches where egg quality was deemed to be good (>30% survival rate at 3 dpf in untreated controls). Each batch included two control groups; uninjected control and 500 ng/µL Cas9 injection control. The survival rate was recorded at 3 dpf and the treatment mortality was calculated against the mortality rate of control using Schneider-Orelli's formula below.

$$\text{Corrected \%} = \left(\frac{\text{Mortality \% in treated group} - \text{Mortality \% in control group}}{100 - \text{Mortality \% in control group}} \right) * 100$$

Larvae from the injected group and controls were sampled between 3 – 5 dpf for subsequent gDNA extraction and assessment of mutation status by molecular analysis. The larvae were euthanised by an overdose of anaesthetic (benzocaine 50-60 ppm,

Sigma-Aldrich, St. Louis, USA) and stored subsequently in 100% ethanol at 4 – 8°C prior to analysis.

4.2.5 Extraction of gDNAs using SSTNE

SSTNE buffer (Appendix, Table S1) was prepared and stored at RT. Each larva sample was placed on a nuclease-free glass plate and the yolk was removed. The excess alcohol was removed before putting the dissected larva into a 1.5 mL eppendorf tube containing 200 µL of SSTNE buffer, 20 µL of 10% SDS (sodium dodecyl sulphate-anionic detergent) and 5 µL of Proteinase K (10 mg/µL). The samples were incubated for 3 hrs to overnight at 55°C until the tissues were fully lysed. After the lysis step, the sample tubes were placed into a hot block set to 70°C for 15 minutes to inactivate Proteinase K. 5 µL of RNase A (2 mg/mL) were added to each tube, mixed by inverting and then incubated at 37°C for 60 minutes. For salt precipitation, 161 µL (0.7 × vol) of 5M NaCl was added into each tube, mixed by inverting and left on ice for 10 minutes. Then, samples were centrifuged at 21,000 xG for 10 minutes to precipitate proteins. 250 µL of each supernatant carrying gDNA was pipetted into a new eppendorf tube. The same volume of isopropanol (250 µL) was added and mixed by 6 sharp (rapid and abrupt) inversions. The samples were left on ice for 5-10 minutes and then centrifuged at 21,000 xG for 10 minutes to pellet gDNA. The supernatant was removed by pipetting avoiding touching the gDNA pellet and briefly spun down to remove excess isopropanol by pipetting. gDNA pellets were washed twice with 75% ethanol. All traces of ethanol were removed by evaporation and the gDNA pellets were resuspended in 10 – 20 µL of 5 mM Tris (pH 8.0) depending on the size of the pellet.

4.2.6 Screening of putative mutants

4.2.6.1 Melt curve analysis

Primer pairs were designed to amplify a DNA fragment including the target area of *piwil2* sgRNA1 and sgRNA2 and used for qPCR melt curve analysis (D'Agostino *et al*., 2016) (Table 4.7, Fig. 4.7).

Table 4.7. Primer list used for identification of *piwil2* mutants using qPCR melt curve analysis

Primer pair	Forward primer (5'-3')	Reverse primer (5'-3')	PCR product size of WT	Annealing temp (°C)
piwil2 sgRNA1	AACAGGTAACTGCTGTCTGCAT	TTGGTTTCTTGCCAGGTTGACTT	210 bp	56.5
piwil2 sgRNA2	TAGGTGAGAATTAGGTGTGGTTT	TGCACAATGCATGAGTCCTAC	211 bp	55.5
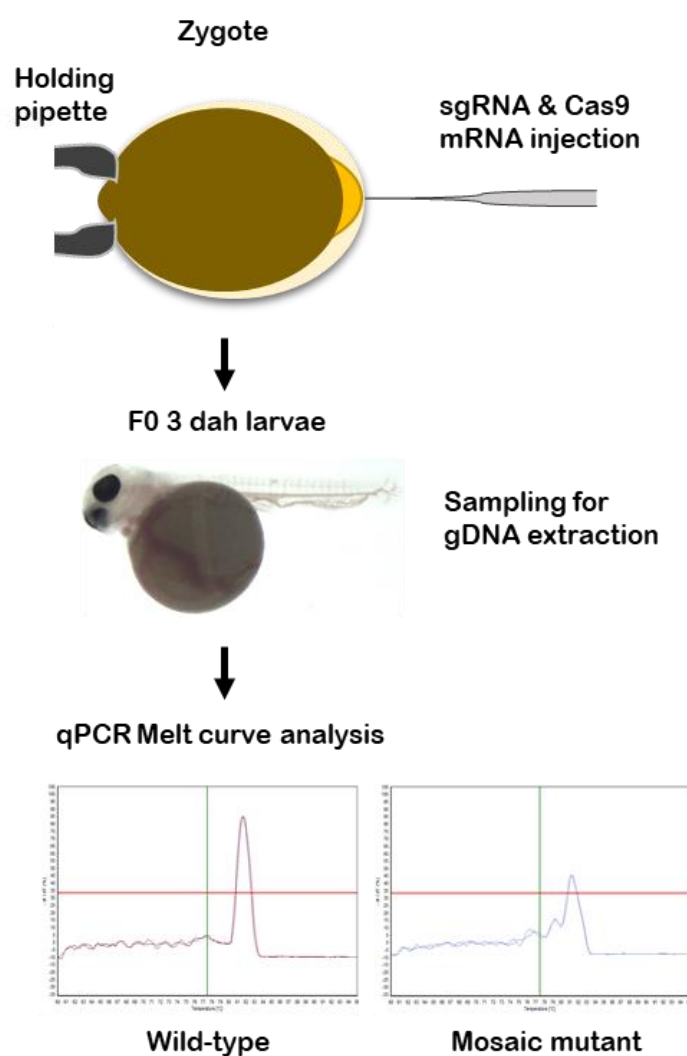
Figure 4.7. Schematic diagrams for identifying mutants induced by *piwil2* CRISPR/Cas9 using melt curve analysis. F0, founder animals (born from microinjection at 1-cell embryos).

Prior to use within the assay, each primer pair was validated by cloning and sequencing of the PCR product from uninjected control gDNA samples ($n = 2$) and 100% sequence identity was confirmed. For cloning and sequencing, pGEM T-easy vector systems (Promega, Madison, USA) and Sanger sequencing (GATC Biotech, Konstanz, Germany) were used, respectively. Each qPCR reaction was of a total volume of 10 µL containing 1 µL of gDNA (70 - 350 ng), 5 µL of SYBR green mix (Luminaris Color HiGreen qPCR master Mix, Thermo Fisher, Waltham, USA), 0.7 µM of each forward and reverse primer and MiliQ water up to 10 µL. Mastercycler® ep realplex (Eppendorf, Hamburg, Germany) was used and the qPCR thermal cycling protocol was: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, 56.5 (sgRNA1) or 55.5°C (sgRNA2) for 15 sec and 72°C for 20 sec. It was followed by melt curve analysis of 95°C for 15 sec, 60°C for 15 sec, a ramp increment at 0.023°C/sec from 60 to 95°C with a continuous fluorescence detection and 95°C for 15 sec. All samples (*piwil2* CRISPR/Cas9 injected larvae, Cas9 injected control, uninjected control) were analysed in duplicate together with non-template controls ($n = 3$). The total number of the injected larvae subjected to mutant screening was 60 (20, 20, 20), 52 (12, 20, 20) and 49 (10, 19, 20) for 100, 150 & 250 ng/µL of sgRNA1 and 59 (20, 20, 19), 50 (19, 11, 20) and 51 (20, 11, 20) for 100, 150 & 250 ng/µL of sgRNA2, respectively (3 batches for each concentration of sgRNA). The number of both Cas9 injected control (0 ng/µL of sgRNA) and uninjected control subjected to mutant screening was 4 per treatment per batch. Presence of mutations in the target site of each sample was assessed by the shape of its derivative of fluorescence with respect to the temperature (dF/dT) dissociation curves using Mastercycler® ep realplex (Eppendorf, Hamburg, Germany) Software.

4.2.6.2 Sanger sequencing of the target sequence

To confirm the sequence in the target region of sgRNA1 and sgRNA2 in Cas9 injected control, uninjected control and *piwil2* CRISPR/Cas9 injected larvae, randomly picked gDNA samples were subjected to PCR, cloning and sequencing. The PCR reaction consisted of a total volume of 20 µL containing 10 µL MyTaq™ Mix (Bioline, London, UK), 1 µL of gDNA (70 - 350 ng), 0.1 µM of each forward and reverse primer (Table 7) and MiliQ water up to 20 µL. The PCR thermal cycling protocol was as follows: 95°C for 1 min, followed by 35 cycles of 95°C for 15 sec, 56.7°C for 15 sec and 72°C for 30

sec and final elongation at 72°C for 2 min. The PCR products were purified, cloned into pGEM T-easy vector (Promega, Madison, USA) and the plasmids were extracted using GenElute^{TM} plasmid miniprep kit (Sigma-Aldrich, St. Louis, USA). Two clones per larva of Cas9 injected control (2 larvae), uninjected control (2 larvae), sgRNA1 (4 larvae) and sgRNA2 (8 larvae) were cloned and sequenced by Sanger sequencing (GATC Biotech, Konstanz, Germany) in both directions to confirm the existence of mutated sequences.

4.2.7 Statistics

Data are presented as mean ± SEM. Statistical analysis was performed using Minitab 17 (Minitab Inc., State College, USA). Data was transformed when necessary to meet the homogeneity of variance and normal distribution. Significant differences in survival rate between uninjected control and phenol red-injected larvae was tested by Two-Sample t-test ($p < 0.05$). The significant difference in data between two sgRNAs and different concentrations of sgRNAs (treatment mortality and mutation rate) were tested by two-way ANOVA. Two-Sample t-test was used to evaluate the significant difference in mutation rates between two sgRNAs at each concentration ($p < 0.05$).

4.3 Results

4.3.1 Phenol red validation

Through testing, it was identified that highly efficient microinjection pipettes must have an opening diameter in the range of 9-12 µm with a sharp tip (Fig. 4.8). In addition, the shape should be rapidly tapered to support the tip well (Fig. 4.8A), otherwise, the long narrow micropipettes (Fig. 4.8B) will be bent or broken, but not penetrating the chorion.

     Survivals in control and microinjected embryos were comparable at hatching (4 dpf) with 76.5 ± 3.9% *vs.* 77.2 ± 4.2% ($p = 0.907$) and at pre-first feeding (7 dpf, 3 dah) with 59.6 ± 2.2% *vs.* 60.5 ± 13.7% ($p = 0.949$), respectively (Table 4.7). The average microinjection speed was 1.24 embryo/min. Therefore, the capacity of this microinjection system for Nile tilapia is around 370 zygotes per batch when performed over a period of 5 hrs by one person. There is a potential to increase the capacity when perfomed by more than two people.
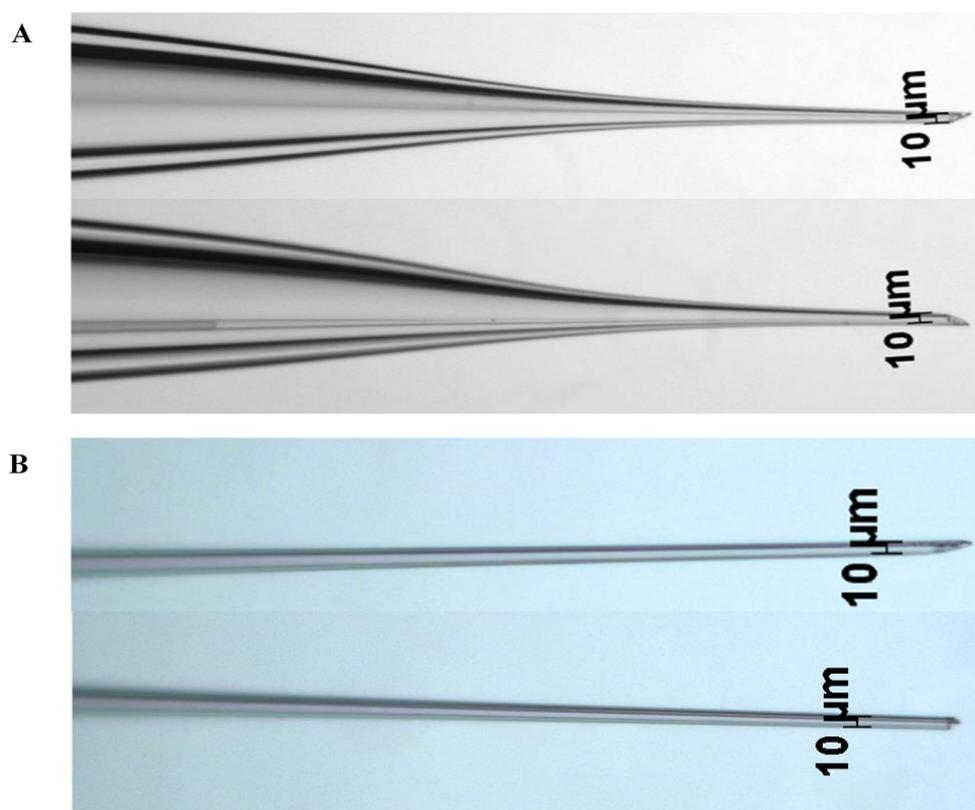
Figure 4.8. Examples of microinjection pipettes for Nile tilapia eggs. (A) Efficient micropipettes. (B) inefficient micropipettes. Scale bars indicate the outer diameter of each tip of micropipette.

Table 4.7. Survival rate of control and phenol red-injected eggs assessed at hatching (4 dpf) and pre-first feeding (7 dpf, 3 dah). Average data were presented as mean ± standard error of the mean (SEM), $n = 3$ independent batches, 80 – 173 larvae per treatment

Egg batch	Stage	Survival rate of control	Survival rate of phenol red-injected eggs
1st batch	Hatching	59/83, 71.1%	61/83, 73.5%
	Pre-first feeding	49/83, 59%	57/83, 68.7%
2nd batch	Hatching	42/50, 84%	58/80, 72.5%
	Pre-first feeding	28/50, 56%	27/80, 33.8%
3rd batch	Hatching	84/113, 74.3%	148/173, 85.5%
	Pre-first feeding	72/113, 63.7%	137/173, 79.2%
Average	Hatching	76.5 ± 3.9%	77.2 ± 4.2%
	Pre-first feeding	59.6 ± 2.2%	60.5 ± 13.7%

4.3.2 PGC labelling using eGFP constructs

There was no apparent specific fluorescent labelling of putative PGCs observed using any of the four mRNA constructs (eGFP-3'URT of *vasa*, *nanos3*, *piwil1* or *piwil2*) as assessed in 20 – 40 embryos at 3, 4 and 5 dpf (pharyngula to hatching stage) in three independent egg batches (Fig. 4.9). The egg yolk was auto-fluorescent and some uneven body parts showed a slight fluorescence, but there were no labelled PGCs with fluorescence.
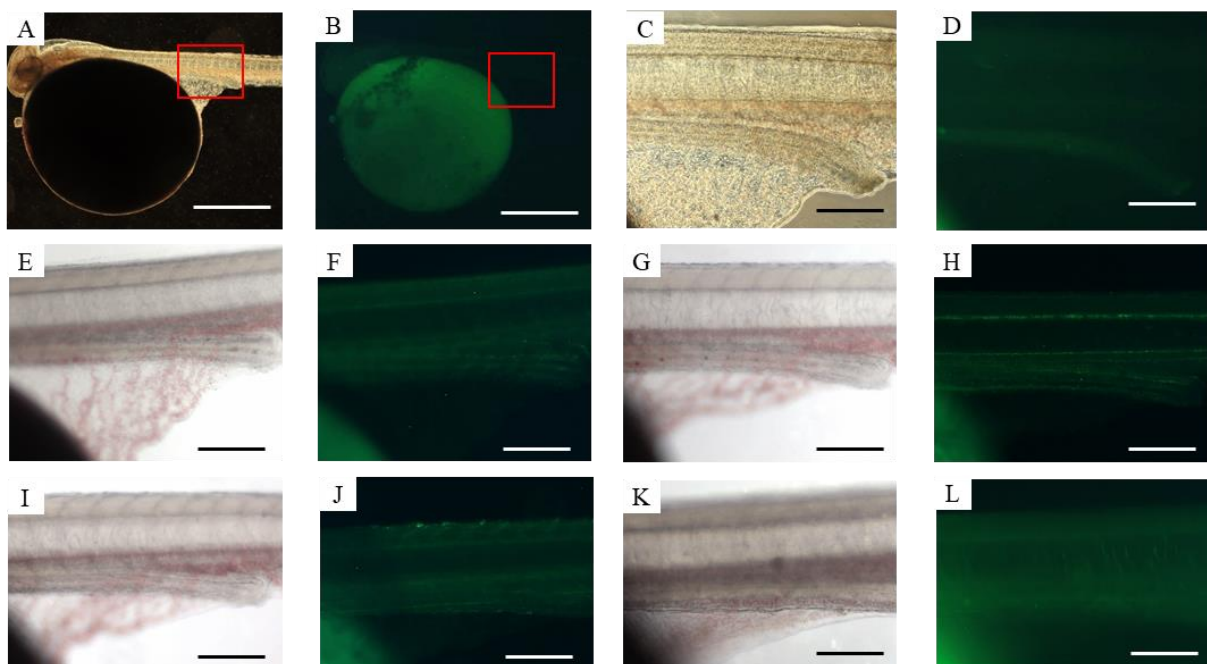
Figure 4.9. Example images of non-fluorescent 4 dpf larvae injected with PGC labelling RNA constructs. (A-D) Control larvae. (A&B) A bright field and FITC filter image of control, respectively, a boxed area indicates the location of gonadal anlagen and this area was enlarged in C-L. (C&D) Enlarged image of control in the bright field and FITC filter, respectively. (E&F) eGFP-3'UTR of *vasa* injected larvae, (G&H) eGFP-3'UTR of *nanos3* injected larvae, (I&J) eGFP-3'UTR of *piwil1* injected larvae and (K&L) eGFP-3'UTR of *piwil2* injected larvae. Scale bars are 1 mm in A-B and 0.2 mm in C-L.

4.3.3 Efficiency of *piwil2* sgRNA1 and sgRNA2 at three different concentrations

The treatment mortalities between *piwil2* sgRNA1 and sgRNA2 ($F = 0.06$, $p = 0.807$), among different concentrations (100, 150 and 250 ng/µL) ($F = 0.08$, $p = 0.926$) as well as between the interaction ($F = 0.44$, $p = 0.655$) were not significantly different (Fig. 4.10A). However, mutation rates of sgRNA2 were significantly higher than sgRNA1 ($F = 270.57$, $p < 0.001$) without significant difference among different concentrations ($F = 0.35$, $p = 0.715$) and the interaction ($F = 0.04$, $p = 0.962$) (Fig. 4.10B). At all three concentrations tested, sgRNA1 showed significantly lower mutation rates than sgRNA2 at 100, 150 and 250 ng/µL (1.7 ± 2.9, 5.0 ± 8.7 and $3.3 \pm 5.8\%$ *vs*. 91.6 ± 10.2, 98.2 ± 3 and $97.0 \pm 5.2\%$, respectively).

Sanger sequencing of the target area in a limited number of randomly selected sgRNA1-injected larvae ($n = 4$) and sgRNA2-mediated mutant larvae ($n = 8$) confirmed the presence of mutated sequences in the injected animals by checking two clones per larva (Fig. 4.11). Seven out of eight sgRNA2-mediated mutant larvae possessed two differently edited sequences from the two clones while one mutant had identical mutated sequence from the two clones (Fig. 4.11).

A

B
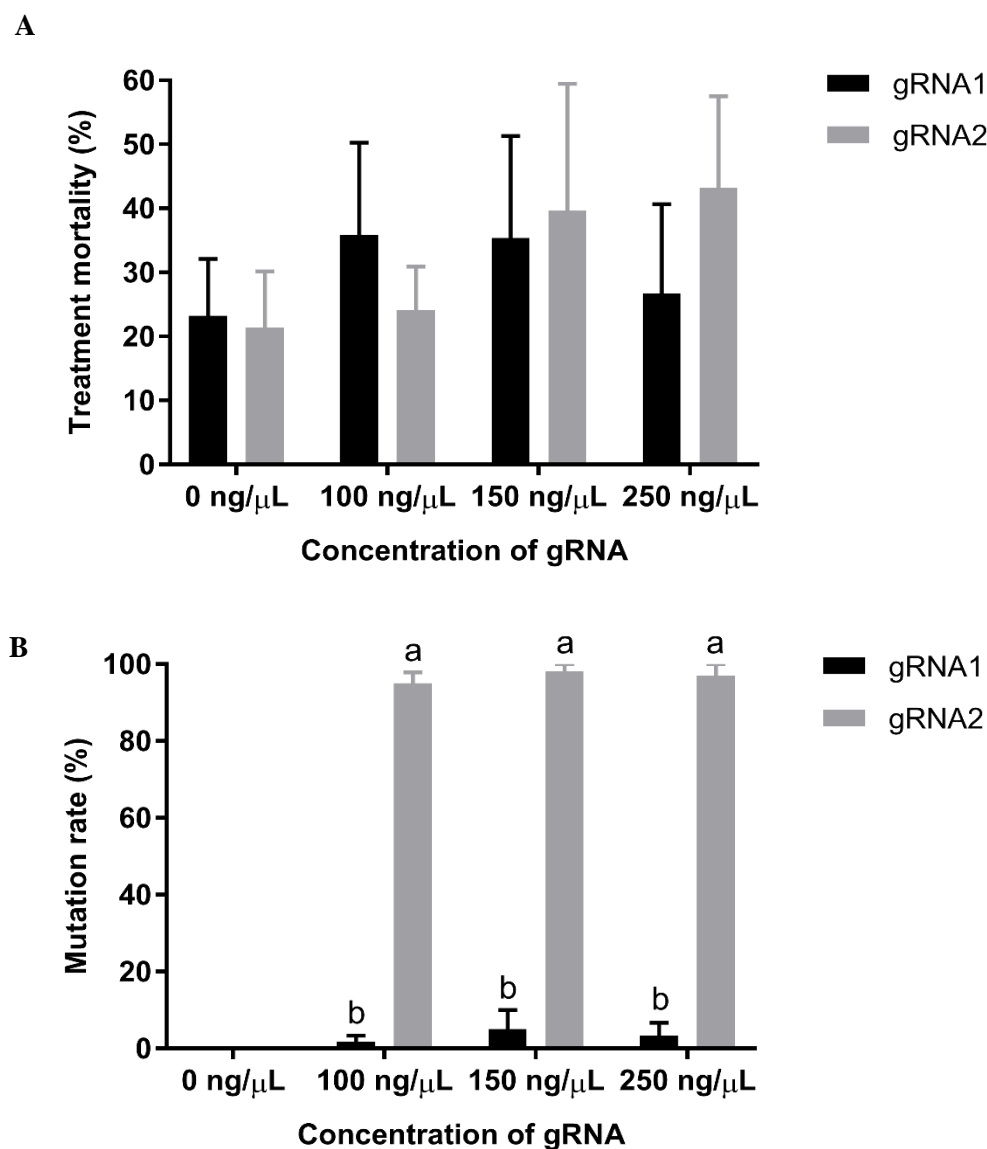
Figure 4.10. Treatment mortality and mutation rates. (A) Treatment mortality and (B) mutation rate induced by different concentrations of *piwil2* sgRNA1 or sgRNA2 together with 500 ng/µL of Cas9 mRNAs. Data were collected from three independent batches and the treatment mortality was recorded at 3 dpf. Putative mutants were screened by qPCR melt curve analysis using individual larva sample (3 – 5 dpf). Cas9 injected control (0 ng/µL of sgRNA) was included ($n = 4$ per batch). Data were presented as mean ± SEM with $n = 3$ batches, 10 to 20 larvae per treatment per batch. Superscripts denote statistically significant difference between sgRNAs at each concentration ($p < 0.05$).

***Piwil2* sgRNA1 target site**

```
AGGTACTGGCCAGACAGGAAAGCCAGTTTGTGGGCATATTTGCAGGGTGCTGGAACACGAATGGTGCCGGGCCAGTTCCAGTACAAATGGCACATCTTGAAAGTCAACCTGGCAAGAAAC  wild type
AGGTACTGGCCAGACAGGAAAGCCAGTTTGTGGGCATATTcGCAGGGTGCTGGAACAC-------------CCAGTTCCAGTACAAATGGCACATCTTGAAAGTCAACCTGGCAAGAAAC  -13, frame-shift
```

***piwil2* sgRNA2 target site**

```
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCTCATTGGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  wild type
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCgtggctTCA--GGGGGTCTGGAAgATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGA  +4, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCTCATccGGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  +1, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCTttccTGGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  +1, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTggCTCAgTGGGGGcCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG   0, in-frame
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCgAACCCTGAAGCCACGGATCAGTTCCTC-cTGGGGGcCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -1, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCT----GGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -4, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCC-------gtggCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -7, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCT-------GTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -8, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTC--------aGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -8, frame-shift
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAGTTCCTC----------TGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -10, frame-shift (*)
TACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCC----------------------TGGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -24, in-frame
TACCTCATAGTACTTCTGCAGTGCAGCC------------------------------------------------AAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -48, in-frame
TACCTCATAGTACTTCTG-----------------------------------------------GGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAGCTGGAAGGGGG  -49, frame-shift
TACCTCATAGTACTTCTGCAG--------------------------------------------------------------------------------GAAGGGGG  -91, frame-shift
```

Figure 4.11. Sequences of *piwil2* mutants edited by CRISPR/Cas9. Each target area of sgRNA1 and sgRNA2 is boxed and PAM sequence is shaded and in bold. Dash means deletion and bold lowercase is inserted nucleotides or substitution. The number on the right indicates the nucleotide length difference and whether it is frame-shift or in-frame mutation. The target site of sgRNA1 ($n = 4$) and sgRNA2 ($n = 8$ mutant larvae) was cloned and sequenced by Sanger sequencing in both direction (two clones per larva). Only 1 out of the 8 clones had a mutated sequence in sgRNA1 samples and 15 out of 16 clones of sgRNA2 were sequenced and showed all had mutated sequences with two clones that were identical (*).

4.4 Discussion

The overarching aim of this study was to optimise the application of type II CRISPR/Cas9 system to edit the *piwil2* gene in Nile tilapia. As part of this process, a number of basic methodological steps had to be developed and optimised in a series to validate a reliable and effective treatment method. At its conclusion, this study demonstrated a highly efficient genome editing system using CRISPR/Cas9 and microinjection system for Nile tilapia which has the capability to treat over 300 embryos per spawning batch with a subsequent larval mutation rate in excess of 90%.

One of the major hurdles in microinjection system in Nile tilapia embryos was the hard chorion of fertilised eggs and it was addressed by developing the holding pipette and the recipe for the stout and sharp microinjection needles suitable for tilapia embryos. Compared to the conventional injection method via the micropyle, it is much faster as it does not require fine positioning of the embryos. In addition, the microinjection procedure set up in this study did not show a detrimental effect on the survival of phenol red-injected tilapia larvae, which is a harmless visual aid used in many teleost species in microinjection procedures (Layden *et al*., 2013). The other restriction was the short time window for microinjection at 1-cell stage in Nile tilapia. To maximise the injection time window, water temperature and fertilisation strategy were manipulated. According to the result of Irawan (1993), high viability in ovulated eggs can be obtained up to 6 hrs post-stripping when the eggs are stored in the incubation water at 25 – 27°C before in vitro fertilisation. The highest viability was obtained from eggs stored for 1 and 2 hrs with 90.6 and 84.7% hatching rates relative to control (immediately fertilised eggs without storage), respectively. Viability gradually reduced to 73.9% and 64.3% hatching rates relative to control when eggs were stored for 4 hrs and 6 hrs, respectively (Irawan, 1993). In addition, when fertilised eggs are kept at 20 – 22°C, which slows down cell division, the fertilized eggs remain at the 1-cell stage for up to 2.5 hrs (Rahman and Maclean, 1992). In this study, the storage of unfertilised eggs and the cold water (20 – 22°C) treatment of 1-cell stage embryos were combined: half of the eggs were fertilised and kept at 20 – 22°C until they were subjected to microinjection while the rest of the unfertilised eggs were kept at 25 – 27°C for up to 2.5 hrs prior to fertilisation. By doing so, the injection time window of 1-cell stage could be extended up to 5 hrs in the current study and species which allowed for an increased number of embryos injected. The average number of injected eggs in the sgRNA optimisation experiments was 372 ± 38

per batch (*n* = 9), while the total number of eggs per batch was 825 ± 77 (*n* = 9). The average injection speed was 1.24 embryo/min, so the capacity of this system is around 370 embryos per batch in a given day when performed by one person. It was noted that the quality of eggs can vary greatly, and poor quality eggs (low fertilisation rate) can impact on the microinjection success. Therefore, any egg batch with less than 30% hatching rate in untreated controls were not included in this study due to the severe loss in sample number at later developmental stages. The average hatching rate of uninjected controls was 65.5 ± 4.9% (*n* = 12). These results clearly indicate the importance of using good quality gametes for microinjection study. Compared to injecting via the micropyle of newly fertilised tilapia embryos which aimed to inject 50 embryos per day (Maclean *et al*., 2002), the microinjection system set up in this study has the capacity to provide feasible number of samples for laboratory-scale experiment. It is a viable research tool to deliver genome editing molecules into zygotes; however, it is not applicable to the large-scale treatment in commercial sector due to the high cost and time. A recent study reported an alternative delivery method of exogenous biomolecules into eggs using liposomes, which is one of the widely used non-viral carriers for biomolecules (Ain *et al*., 2015). Kumari *et al*. (2017) demonstrated a successful delivery of a lipophilic marker, Rhodamine B isothiocyanate, into the egg yolk by bath immersion of unfertilised eggs in Atlantic salmon. The development of an efficient, safe, simple and cost-effective gene delivery method such as a liposome-mediated strategy or any other nanodelivery method would enable large-scale commercial application. Further research and development on novel delivery methods is required.

The ultimate goal of this research was to explore the use of genome editing to affect the development of the PGCs which has the potential to subsequently render the animal sterile. It takes 6 months for tilapia to develop reproductive competence, thus tools that could indicate PGC fate and the gonadal development earlier in the ontogeny of animals would be useful to fast-track analysis of gene editing effect. To this end, development of four GFP linked reporter constructs was attempted, whereby the 3'UTR of *vasa, nanos3, piwil1* & *piwil2* were conjugated to ORF of the reporter gene, eGFP. These targets were chosen because 3'UTR of *vasa* and *nanos3* are known to have gene regulatory regions which stabilise the mRNAs in germline cells (Knaut *et al*., 2002; Škugor *et al*., 2014a; Wolke *et al*., 2002). For example, in zebrafish, *nanos3* is targeted by microRNA-430, causing degradation in soma while germline specific RNA binding proteins present in germline cells such as Dnd bind to the 3'UTR of *nanos*3 to protect the transcripts from

degradation by microRNA-430 (Kedde *et al*., 2007; Mishima *et al*., 2006; Slanchev *et al*., 2009). It is unknown whether the 3'UTRs of *piwil1* or *piwil2* serve a similar function; therefore, their 3'UTRs were included to test the function in PGC specific localisation. It was aimed to develop functional PGC-labelling constructs which can serve as an indicator of physiological impacts of knockout or knockdown of the target gene on PGCs (Gross-Thebing *et al*., 2017; Li *et al*., 2014). However, none of the constructs tested in this study appeared to label PGCs. Although only a single concentration (350 nM, 132 – 176 ng/µL) of the PGC labelling constructs was tested, it was selected based on the concentration used in previous studies: in Nile tilapia, 50 ng/µL of eGFP-*vasa* 3'UTR RNA was injected (Li *et al*., 2014), 150 ng/µL of GFP-*nanos3* 3'UTR in grouper (Sun *et al*., 2017), 300 ng/µL of GFP-*nanos3* 3'UTR in barfin flounder (*Verasper moseri*) (Goto *et al*., 2015) and 200 ng/µL of GFP-*vasa* 3'UTR in rainbow trout, zebrafish and Nibe croaker (*Nibea mitsukurii*) (Yoshizaki *et al*., 2005). The possible cause for the failure might be the reduced sensitivity of the fluorescent microscope used in this study since the autofluorescence of egg yolk was barely detected while the egg yolk showed bright autofluorescence in the result of Li *et al*. (2014). The other potential explanation for the lack of PGCs labelling using eGFP-3'UTR constructs might be the rapid deadenylation of the RNA constructs in the embryos. It may need a 5' cap as shown in zebrafish where $5'm^7G$ (7-methylguanosine) capped GFP was expressed more strongly than A-capped GFP injected larvae (Mishima *et al*., 2006). This would need further confirmation. Even though PGC-labelling molecules are unavailable, the initial purpose to explore the physiological impact of gene editing on PGCs in the subsequent investigation can be achieved through histological observation of PGCs between hatching and early larval stage (pre-first feeding) in Nile tilapia, identified by their location and morphological features (Ijiri *et al*., 2008).

For a successful gene knockout, optimisation of the components of CRISPR/Cas9 is required. In Nile tilapia, various ratios of sgRNA to Cas9 mRNAs have been used from 50 ng/µL of sgRNA and 500 ng/µL of Cas9 mRNA (1:10) to 250 and 500 ng/µL (1:2) (Chen *et al*., 2017; Feng *et al*., 2015; Jiang *et al*., 2016; Li *et al*., 2015; Zhang *et al*., 2014b). Li *et al*. (2014) tested a constant sgRNA concentration (50 ng/µL) with increasing Cas9 mRNA concentrations (100, 300 & 500 ng/µL) targeting *nanos2* and *nanos3* in Nile tilapia, resulting in Cas9 dose-dependent mutation and mortality. With the decrease in Cas9 mRNA concentration, the survival rate following injection increased from 12.6 to 33% in *nanos2* and 7 to 27% in *nanos3*, while the proportion of indel

mutation rate decreased from 51 to 13% in *nanos2* and 38 to 8% in *nanos3* (Li *et al.*, 2014), but none of the tested concentrations showed an efficient mutation rate or appropriate survival rates. Therefore, in this study, we tested three different concentrations of two sgRNAs (100, 150 & 500 ng/µL) with a single concentration of Cas9 mRNA (500 ng/µL) in triplicates to determine the optimal concentrations for high mutation rates and low mortality in this species. However, there were no significant differences in mutation efficiency and mortality in relation to sgRNA concentrations. In zebrafish, increasing ratio of sgRNAs to a constant Cas9 mRNA amount resulted in sgRNA dose-dependent gene editing efficiency (Jao *et al.*, 2013). This discrepancy might be caused by the different range of the tested ratios. In this study, three different ratios of sgRNA to Cas9 mRNAs were 1:5, 1:3.3 and 1:2 which were above the range tested in zebrafish (*i.e.* 1:150, 1:25 and 1:5), suggesting that all three ratios tested in this study may, in essence, be at a saturated level, thus maximum potential mutation rate was observed.

Piwi family has two distinct domains: PAZ domain, an RNA binding motif, and Piwi domain, a similar structure of RNase H catalytic domain (Simon *et al.*, 2011). The Piwi domain is known to act as a catalytic engine in RISC for RNA interference (Liu *et al.*, 2004). Thus, both sgRNAs were designed to target a conserved amino acid sequence within Piwi domain of *piwil2*. Interestingly, there was a significant difference in mutation rate between *piwil2* sgRNA1 and sgRNA2 in this study. There are some possible reasons for these results. One explanation could be an inappropriate synthesis of sgRNA. In this study, both sgRNA1 and sgRNA2 were purified after synthesis and the right size and integrity were confirmed by running on a gel with RNA size marker thus it is presumed this was not a cause of the observed differences. Another explanation could be a difference in intrinsic properties of sgRNA which can impact on the cleavage efficiency such as RNA secondary structures (Ma *et al.*, 2015) and its affinity to Cas9, *e.g.* preferentially binds to sgRNAs containing purines in the last 4 nucleotides of the spacer sequence (Wang *et al.*, 2014b). In this study, there was no significant difference in predicted secondary structures between the two sgRNAs. In addition, both sgRNAs used in this study were designed by sgRNA designing tools which use their own algorithm and formula to find highly efficient sgRNAs (Bae *et al.*, 2014; Doench *et al.*, 2014). The dramatic difference in mutation efficacy of the two sgRNAs implies that there is an obvious caveat in current sgRNA designing tools as the criteria for computational prediction of sgRNA efficiency are derived from limited data (Carrington *et al.*, 2015).

Further studies are needed to determine the underlying mechanisms governing target site recognition and interaction between sgRNA and Cas9 (Jao *et al*., 2013). Finally, it must be acknowledged that the intrinsic properties of the target sites in terms of chromatin/epigenetic states, transcription activities and high GC-content, could be the main drivers behind the variable mutation rates observed (Doench *et al*., 2016; Jao *et al*., 2013; Wu *et al*., 2014). Without clear indication as to the physical and/or biochemical properties that could influence the CRISPR editing efficiency, the pragmatic approach to future sgRNA design would be to test multiple target specific sgRNAs.

One of the main concerns of CRISPR/Cas9-mediated genome editing is the off-target effects (Jiang and Doudna, 2017). The match of 12 nt seed sequence is known to be important for the binding of Cas9/sgRNA complex to DNA (Cong *et al*., 2013; Jinek *et al*., 2012). In this study, *in silico* off-target analysis confirmed that there is no site which has an identical 12 nt seed sequence adjacent to PAM for both sgRNA1 and 2. In addition, the analysis of bulge type off-targets (1 bp bulge with 2 mismatches) confirmed no bulge type off-target for sgRNA1 and four for sgRNA2 with one present in an intron and the rest of three not located in genes, which is unlikely to have an impact on the function of the off-target. However, it is still recommended to conduct unbiased genome-wide screening such as BLESS (direct *in situ* breaks labelling, enrichment on streptavidin, and next-generation sequencing), GUIDE-seq (genome-wide, unbiased identification of DSBs enabled by sequencing) and HTGTS (high-throughput genome-wide translocation sequencing) to detect unintended changes, even though there is a low chance of off-target activity (reviewed by Zischewski *et al*., 2017). This was attempted in the present work but unaccomplished due to time constraints.

Melt curve analysis was chosen to screen mutants produced by CRISPR/Cas9 in this study, as it can detect the formation of heteroduplex which implies the heterozygous mutation (mono- or biallelic) (D'Agostino *et al*., 2016). It was reported that it has a detection limit with indel less than 15 bp which are not clearly distinguishable from the derivative melt curves (D'Agostino *et al*., 2016); however, it is a cheap and fast screening method that can be performed at a high throughput. The presence of mosaicism shown in F0 animals in the melt curve analysis was confirmed by Sanger sequencing of the target area. The high mosaicism in F0 animals mediated by CRISPR/Cas9 system was reported in previous studies (Edvardsen *et al*., 2014; Wargelius *et al*., 2016; Zhong *et al*., 2016). Although melt curve analysis indicates the existence of mutation and mosaicism, it provides no further information on genotypes of the F0 animals generated by

CRISPR/Cas9. The following step is, therefore, to characterise the genotypic features of the mutants as suggested by Mianné *et al*. (2017) as the outcome of CRISPR/Cas9 genome editing can be unpredictable.

In conclusion, this study demonstrated a treatment that reproducibly results in highly efficient mutation rate (> 90%) without significant treatment mortality. This technique can, therefore, be used to investigate the function of *piwil2* on the survival of PGCs in Nile tilapia with the long-term goal to explore its potential as a means to induce sterility in the species.

CHAPTER 5

PHYSIOLOGICAL IMPACT OF *PIWIL2* KNOCKOUT ON PGCS IN NILE TILAPIA INDUCED BY CRISPR/CAS9

5.1 Introduction

To seek an alternative sterilisation method in fish, several genes known to play important roles in PGCs maintenance and survival and apoptosis in gonads were screened in Nile tilapia and *piwil2* gene was selected as a target for gene editing and ablation of PGCs in Nile tilapia (Chapter 2 and Chapter 3). Thereafter, a microinjection system and CRISPR/Cas9 system targeting *piwil2* gene were optimised in Nile tilapia embryos to obtain high mutation rate and low mortality (Chapter 4). In the current Chapter, the CRISPR/Cas9 gene editing system was tested and the phenotypes of the resulting tilapia larvae were screened to determine the physiological impact of *piwil2* KO on PGC development to interpret the capability of this method to induce sterility in Nile tilapia.

piwil genes encode germline specific Argonautes, silencing mobile genetic elements to maintain germ cell fate (Ewen-Campen *et al*., 2010; Houwing *et al*., 2007). Argonautes including the Piwi family have two conserved domains: (1) PAZ domain, an RNA binding motif and (2) Piwi domain, similar in structure to the RNase H catalytic domain (Simon *et al*., 2011). The Piwi domain acts as a catalytic engine in RISC for RNA interference (Liu *et al*., 2004). In this study, sgRNA was designed to guide Cas9 enzyme to cleave the Piwi domain of the *piwil2* gene in Nile tilapia. Previous *piwi* loss of function studies have shown a severe deficiency in germ cell development and spermatogenesis in various species including *Drosophila* (Cox *et al*., 2000; Lin and Spradling, 1997), mice (Carmell *et al*., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa *et al*., 2004) and zebrafish (Houwing *et al*., 2007, 2008).

As the end result of the genome engineering method is intended to be sterility, functional studies can only rely on the screening of injected founder animals as the production of stable homozygous mutant lines is not viable. One of the significant advantages of CRISPR/Cas9 system is the high efficiency of genome editing which can produce biallelic mutants in founder (F0) individuals, allowing a direct phenotypic analysis without the need to generate F1 and F2 generations (Jao *et al*., 2013). This method can, therefore, be used to investigate gene function in species which have a long generation time such as salmonid or target genes with the aim to cause sterility.

The physiological impact of gene KO on PGCs in fish can be observed through the development and application of PGC labelling constructs such as green fluorescent protein (GFP) conjugated with 3'UTR regions of PGC specific gene such as *vasa* or *nanos3* (Gross-Thebing *et al*., 2017; Li *et al*., 2014). Initially, it was intended to develop

a PGC labelling construct which can be co-injected with CRISPR/Cas9 into the zygote to visualise the presence of PGCs during early ontogeny. However, these labelling constructs were unable to label PGCs and thus could not be applied in subsequent research (Chapter 4). An alternative means to confirm and monitor PGC development would be through histological observation (Ijiri *et al.*, 2008). At early developmental stages prior to first feeding in larvae, PGCs can be found in gonadal anlagen in Nile tilapia and they can be identified on the histologcal sections based on location and morphological features such as large round-shaped cells with large nucleus (Ijiri *et al.*, 2008).

The aim of this study was to investigate the function of the *piwil2* gene using the optimised CRISPR/Cas9 system developed in Chapter 4. The mutants were screened by melt curve analysis and the physiological impact of *piwil2* KO on PGCs was investigated by histological observation of gonadal anlagen at early developmental stages in Nile tilapia larvae.

5.2 Materials and Methods

5.2.1 *piwil2* CRISPR/Cas9 and microinjection

5.2.1.1 Preparation of *piwil2* sgRNA2 and Cas9 mRNAs

gRNA and Cas9 mRNAs were prepared as described in Chapter 4. Briefly, sgRNA2 was designed in exon 21 which is the conserved domain of PIWI and checked not to have potential off-target sites (no more than 12 nt is identical out of 20 nt sgRNA sequence together with PAM) using RGEN tool and NCBI BLAST. Both sgRNA scaffold and Cas9 ORF sequences were obtained from pT7-gRNA and pT3TS-nCas9n plasmids, respectively, supplied by Wenbiao Chen (Addgene plasmid #46759 and #46757, respectively) (Jao *et al.*, 2013). A PCR approach was taken for sgRNA template production as described in Chapter 4 and according to Shao *et al.* (2014). The Cas9 sequence was type II CRISPR RNA-guided endonuclease Cas9 (*S. pyogenes*) which cause DSBs at the target site. The Cas9 ORF template was amplified by PCR and used as template DNA for *in vitro* transcription using mMessage mMachine T3 kit (Thermo Fisher, Waltham, USA) and the resultant capped Cas9 mRNA was purified using RNeasy

Mini Kit (Qiagen, Hilden, Germany). The sgRNA2 and the capped Cas9 mRNA solution was diluted and aliquoted for single use and stored at -20°C until use.

5.2.1.2 Injection of *piwil2* sgRNA2 and Cas9 mRNAs and sampling

Based upon the optimisation result in Chapter 4, 150 ng/µl of *piwil2* sgRNA2 was injected with 500 ng/µl of Cas9 mRNAs into 1-cell stage embryos using two egg batches produced from two females (XX) crossed with a single supermale (YY) (Scott *et al.*, 1989). The tilapia gametes were sourced from the Tropical Aquarium, Institute of Aquaculture, University of Stirling. Each batch included a uninjected control group. The total number of injected embryos was 251 and 365 and the control was 367 and 267 for each batch. All experimental embryos/larvae were maintained in tanks designed for prevention of escapee as approved by the GMSC of the University of Stirling (Risk assessment US-GM042-IOA-AD). The survival rates were recorded at 3 days after fertilisation (dpf) and 3 days after hatch (dah). The controls and injected group larvae were sampled at 3 dah for subsequent molecular and histological analyses. Fish were anaesthetised with benzocaine (Sigma-Aldrich, St. Louis, USA) and dissected; each head and tail were stored in 100% ethanol at 4 – 8°C and the body trunk was fixed in 4% PFA for 24 hours at 4°C and transferred to 70% EtOH in DEPC treated water and stored at -20°C until further analysis.

5.2.2 Initial mutant screening

5.2.2.1 Extraction of gDNA using SSTNE

SSTNE buffer (Appendix, Table S1) was prepared and stored at room temperature. The excess alcohol was removed before putting the head and tail samples into 1.5 mL eppendorf tubes containing 200 µL of SSTNE buffer, 20 µL of 10% SDS and 5 µL of Proteinase K (10 mg/µL). The samples were incubated overnight at 55°C to fully lyse the tissues. After the lysis step, the extraction procedure was as described in Chapter 4. The gDNA pellets were resuspended in 15-20 µL of 5 mM Tris (pH 8.0) depending on the size of pellet.

5.2.2.2 Screening of putative mutants using melt curve analysis

To amplify DNA template including target area of *piwil2* sgRNA2 the *piwil2* sgRNA2 primer pair (forward, 5'-TAGGTGAGAATTAGGTGTGGTTT-3'; reverse, 5'-TGCACAATGCATGAGTCCTAC-3') were designed and validated as reported in Chapter 4. Each qPCR reaction was of a total volume of 10 µl containing 1 µL of gDNA (70-350 ng), 5 µL of SYBR green mix (Luminaris Color HiGreen qPCR master Mix, Thermo Fisher, Waltham, USA), 0.7 µM of each forward and reverse primer and MilliQ water (EMD Millipore, Burlington, USA) up to 10 µL. Mastercycler® ep realplex (Eppendorf, Hamburg, Germany) was used and the qPCR thermal cycling protocol was: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, 55.5°C for 15 sec and 72°C for 20 sec. It was followed by melt curve analysis of 95°C for 15 sec, 60°C for 15 sec, a ramp increment at 0.023°C/sec from 60 to 95°C with a continuous fluorescence detection and 95°C for 15 sec. All samples were analysed in duplicate together with non-template controls. The presence of mutations in the target site of each sample was assessed by the shape of the first derivatives of the melt curves using Mastercycler® ep realplex (Eppendorf, Hamburg, Germany) Software.

5.2.3 Histological observation of PGCs in *piwil2* mutant individual

5.2.3.1 Tissue processing and paraffin embedding

The fixed tissues were processed and embedded in RNase-free conditions. All glassware, forceps, wax moulds and cassettes used in this process were nuclease-free and ethanol series were prepared with DEPC (diethylpyrocarbonate) treated water (Thermo Fisher, Waltham, USA). The fixed tissues stored in 70% EtOH at -20°C were individually wrapped in histology paper and labelled. Every six wrapped tissues was cassetted together and then dehydrated in an ascending ethanol series, cleared in xylene and then infiltrated with paraffin wax as shown in Table 5.1. Every trunk tissue was embedded in paraffin with an anterior-posterior presentation for transverse section. 2-3 body trunks were embedded together in the same block using a histoembedder (Leica, Wetzlar, Germany). The wax blocks were trimmed and sectioned using a Rotary microtome (Leica, Wetzlar, Germany).

Table 5.1. Tisssue processing procedure for both ISH and H&E samples

Step	Solution	Time (min)	Temp (°C)	Repeat	Condition	Purpose
1	70% ethanol	45	18-22	X 1	Gentle shaking	Dehydration
2	80% ethanol	45	18-22	X 1		
3	90% ethanol	45	18-22	X 1		
4	99-100% ethanol	45	18-22	X 3		
5	Xylene	45	18-22	X 2	Fume hood	Clearing
6	Paraffin wax	60	56-58	X 2	Incubator	Wax impregnation

5.2.3.2 H & E staining

As a preliminary test, the body trunks of 3 dah control tilapia larvae ($n = 5$) were analysed to validate the histological screening method of PGCs for *piwil2* mutant animals. The embedded trunk tissues of control (approx. 2.3 mm in length, Fig. 1) in paraffin were trimmed using a microtome and approximately 300-350 µm of the anterior body trunk was discarded. Then, the first six consecutive serial sections of 5 µm thickness were collected in every 150 µm interval (30 sections) through the whole trunk (Fig. 1). The sections were stained with H&E (Appendix, Table S2) for histological examination under a light microscope (BX51, Olympus, Tokyo, Japan). Finally, the location of PGCs was confirmed in 3 dah control larvae using the interval serial sectioning. The result showed that PGCs are located in the parietal peritoneum (the outer layer attached to the abdominal wall) beneath the nephric ducts in two gonadal anlagen (Fig. 5.1A-D). They constantly appeared from 0.9 mm to 1.4 mm from the rostral side of the decapitated body trunk, which is indicated by a shaded box in Fig. 5.1. It was no longer present at the end of the body cavity (Fig. 5.1E). Based on the preliminary screening result to locate precisely the gonadal anlagen in 3 dah control larvae, the histological analysis of the *piwil2* sgRNA2/Cas9 injected larvae was done on the region A-D as shown in Fig. 5.1. The first 600-650 µm of each trunk was trimmed and then six consecutive serial sections at 5 µm thickness were collected at every 150 µm interval until the end of the body cavity. Histological slides collected in the area from A to D were analysed for every sample and the presence of PGCs with its location and appearance was recorded.
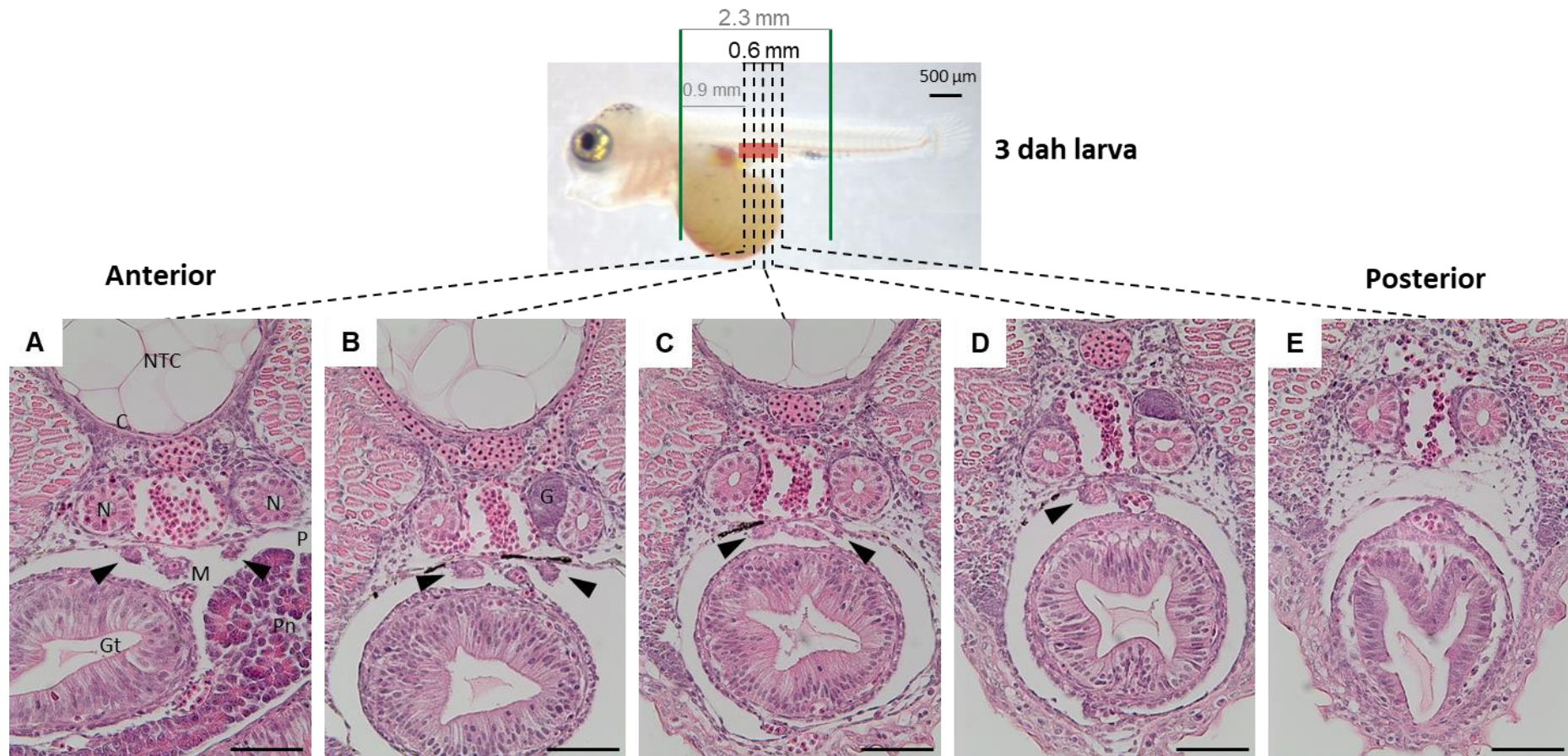
Figure 5.1. Histological observation of gonadal anlagen and PGCs in serial transverse sections of 3 dah control larva. (A-E) Transverse view of the body cavity (anterior-posterior) stained with H&E. The position of each histology section corresponding to images A to E are illustrated by the black dotted lines and the interval between the dotted lines is 150 µm. The embedded trunk tissues of 3 dah control larvae were approx. 2.3 mm in length (indicated by two green lines in the lateral view of 3 dah larva). The red shaded box indicates the location of putative gonadal anlagen. NTC, remnants of the notochord; C, centrum; N, pronephric tubules; G, glomeruli; P, parietal peritoneum; M, mesenteries; Pn, pancreas; Gt, gut. Arrowheads indicate gonadal anlagen. Scale bar = 50 µm (A-E).

5.2.4 PGC localisation through *in situ* hybridisation

To confirm the histological examination of PGCs, an *in situ* hybridisation (ISH) approach was developed to visualise PGCs using germ cell specific *vasa* probes.

5.2.4.1 Preparation of digoxigenin (DIG) labelled *vasa* riboprobes

vasa probe was designed on the highly conserved domain, DEAD box helicase, using NCBI primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/), which is consensus sequences among a number of *vasa* genes in Nile tilapia (GenBank accession No. NM_001279610.1, AB032467.1, XM_019351269 – 78, XM019351972 – 5, XM019351977 – 9 and XM_019358458 – 61). To obtain the *vasa* probe sequence in Nile tilapia, total RNA were extracted from testis and ovary from adult Nile tilapia using the routine TRI reagent (Sigma-Aldrich, St. Louis, USA) extraction method, following mechanical disruption (Mini-BeadBeater, BioSpec Products, Bartlesville, USA) of the samples with 3 mm glass beads (EMD Millipore, Burlington, USA). RNA integrity and yield were confirmed by gel electrophoresis and spectrophotometry, respectively, and then all RNA samples were treated with DNase I (DNA-free DNA Removal kit, Thermo Fisher, Waltham, USA) to remove residual gDNA contamination. cDNA was reverse transcribed from 400 ng of DNase I treated RNA (High capacity cDNA reverse transcription kit, Applied Biosystems, Foster City, USA) using a blend of random hexamer & anchored oligo dT primer at a 3:1 ratio. The target sequence was amplified by routine PCR where each reaction consisted of a total volume of 25 μL containing 12.5 μL MyTaq mix (Bioline, London, UK), 1 μL of cDNA (undiluted), 0.2 μM of each forward and reverse primer (forward, 5'- TTCCCACTATGAGTCCGGTGT -3'; reverse, 5'- CTGCTCCCTCTTTGCGAACT -3') and MilliQ water (EMD Millipore, Burlington, USA) up to 25 μL. The PCR thermal cycling protocol was as follows: 95°C for 1 min, followed by 35 cycles of 95°C for 15 sec, 57.6°C for 20 sec and 72°C for 1 min and final elongation at 72°C for 2 min. The presence of the target was assessed by gel electrophoresis and PCR products were purified using GeneJET PCR purification kit (Thermo Fisher, Waltham, USA). The purified amplicons were ligated to pGEM T-easy vector (Promega, Madison, USA) and transformed into JM109 competent cells (Promega, Madison, USA). The target sequence (814 bp) and the orientation in the extracted plasmid were confirmed by Sanger sequencing (GATC Biotech, Konstanz, Germany) and *vasa* sequences (814 bp) from both male and female were identical. Then, the

plasmid was linearised using appropriate restriction enzymes to be used as a template DNA in *in vitro* transcription to produce antisense and sense riboprobes using DIG RNA labelling kit (Roche, Basel, Switzerland). The integrity of the transcribed riboprobes was checked by electrophoresis on a 1% agarose gel stained with EtBr. The concentration of each riboprobe was quantified by spectrophotometry (Nanodrop, Thermo Fisher, Waltham, USA) and stored at -20°C prior to analysis.

5.2.4.2 ISH in mature testis and 3 dah larvae

The confirmation of functionality of ISH constructs was performed in two iterative steps, firstly in mature testis where *vasa* expression is known to be significant (Kobayashi *et al*., 2000, 2002). Thereafter, it was tested in 3 dah larvae to be reflective of subsequent sample analysis. Mature male tilapia ($n = 2$) were killed by an approved Schedule 1 method (overdose of Benzocaine and severing of spinal column) and dissected to obtain testis. A half of the testis tissue was fixed in 4% PFA for 24 hours at 4°C and the other half was treated with RNAlater for 24 hours at 4°C. The fixed testis tissues were processed and embedded as described above. ISH was performed in the mature testis sample using DIG labelled *vasa* riboprobes, anti-DIG antibody and HNPP detection kit (Roche) in the Laboratory of Prof. José A. Muñoz-Cueto in the Universidad de Cádiz with the support from Dr. Jose Antonio Paullada Salmerón. Briefly, the paraffin-blocks of 3 dah body trunks were sectioned in gonadal anlagen regions. The obtained sections were de-waxed in Xylene, rehydrated in an alcohol and PBS series, post-fixed in 4% PFA and treated with proteinase K (2 µg/mL, 5 min at 37°C). Following the proteinase K treatment, the samples were hybridised (16 h at 55°C) with the pre-denatured probe at 75°C for 5 min. After stringent washing (55°C) and blocking, the slides were incubated with anti-DIG antibody (1/2000, 16 h at 4°C). The slides were then, revealed by incubating with HNPP/Fast Red solution in a dark humified box for 5-7 hours at room temperature. Following a washing step, the slides were mounted with Vectashield antifade mounting medium with DAPI (Vector Lab, Burlingame, USA) and examined by fluorescent microscopy. Then, the same procedure was applied to the 3 dah control larvae samples.

5.2.4.3 Expression level of *vasa* in mature testis and 3 dah larvae

Expression levels of *vasa* in samples used for ISH test were measured by qPCR in order to comment on the signal to noise ratio in the ISH samples using the method as follows. Total RNA were extracted from the mature testis ($n = 2$) and the body trunk of control 3 dah larvae ($n = 4$) using the routine TRI reagent (Sigma-Aldrich, St. Louis, USA) extraction method, following mechanical disruption (Mini-BeadBeater, BioSpec Products, Bartlesville, USA) of the samples with 3 mm glass beads (EMD Millipore, Burlington, USA). The RNA integrity and yield were confirmed by gel electrophoresis and spectrophotometry, respectively, and then all RNA samples were treated with DNase I (DNA-free DNA Removal kit, Thermo Fisher, Waltham, USA) to remove residual gDNA contamination. cDNA was reverse transcribed from 1 µg of DNase I treated RNA (High capacity cDNA reverse transcription kit, Applied Biosystems, Foster City, USA) using a blend of random hexamer & anchored oligo dT primer at a 3:1 ratio. Once produced, cDNA samples were stored at -20°C prior to analysis.

The qPCR primer pair for Nile tilapia *vasa* and *ß-actin* were designed as outlined in Chapter 2. The serial standard samples for each target genes which were previously validated (Chapter 2) were used to calculate absolute copy number. Each qPCR reaction was of a total volume of 10 µL containing 1 µL of cDNA, 5 µL of SYBR green mix (Luminaris Color HiGreen qPCR master Mix, Thermo Fisher, Waltham, USA), 0.3 µM of each forward and reverse primer and MilliQ (EMD Millipore, Burlington, USA) water up to 10 µL. The qPCR thermal cycling protocol was: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, annealing for 30 sec at temperatures optimised for specific primers (Chapter 2, Table 2.2) and 72°C for 30 sec, and completed with melt curve analysis. All samples were analysed in duplicate together with non-template controls and standard samples. The qPCR efficiency of *ß-actin* and *vasa* was 101.1 and 94.5%, and the linearity (R^2) of standard curves was 0.994 and 0.999, respectively. The standard curves ranges used for *ß-actin* and *vasa* were $10^2 - 10^7$ and $10^1 - 10^6$ copies per reaction, respectively. A melt curve analysis confirmed that a single product was generated without primer dimer artefacts. The absolute copy numbers of each target gene were calculated based on its standard curve which was thereafter normalised to *ß-actin* concentration to calculate the relative copy number.

5.2.5 Statistical analyses

The statistical difference in survival rate between control and injected larvae and the relative number of *vasa* transcripts between the testis and 3 dah larvae was determined by unpaired t-test with Minitab 17 (Minitab Inc., State College, USA) ($p < 0.05$). Data are presented as mean ± SD.

5.3 Results

5.3.1 Survival and mutation rate

There was no significant difference in survival rates between control and *piwil2* sgRNA2/Cas9 injected eggs at hatching (3 dpf) and pre-first feeding stage (6 dpf) ($p =$ 0.51 and 0.79, respectively) (Table 5.2). The mutation rates checked at 6 dpf (3 dah) were 98.9% and 92.8% in 1st and 2nd batch, respectively (Table 5.2). There was no apparent difference in the appearance of the live larvae at 3 dah between uninjected control and *piwil2* sgRNA/Cas9 injected larvae (Fig. 5.2).

Table 5.2. Survival and mutation rates of control and *piwil2* sgRNA2/Cas9 injected eggs assessed at hatching (3 dpf) and pre-first feeding (6 dpf, 3 dah). Two independent egg batches were injected with 150 ng/µL of *piwil2* sgRNA2 and 500 ng/µL of Cas9 RNA

Egg batch	Stage	Survival		Mutation rate
		Control	***piwil2* sgRNA2/Cas9**	
1st batch	Hatching	210/367, 57.2%	149/251, 59.4%	
	Pre-first feeding	184/367, 50.1%	128/251, 51.0%	86/87, 98.9%
2nd batch	Hatching	210/267, 78.7%	217/365, 59.5%	
	Pre-first feeding	93/267, 34.8%	90/365, 24.7%	77/83, 92.8%
Average	Hatching	67.9 ± 15.2%	59.4 ± 0.1%	
	Pre-first feeding	42.5 ± 10.8%	37.8 ± 18.6%	95.8 ± 4.3%

(A) Control 3 dah larvae   (B) Injected 3 dah larvae
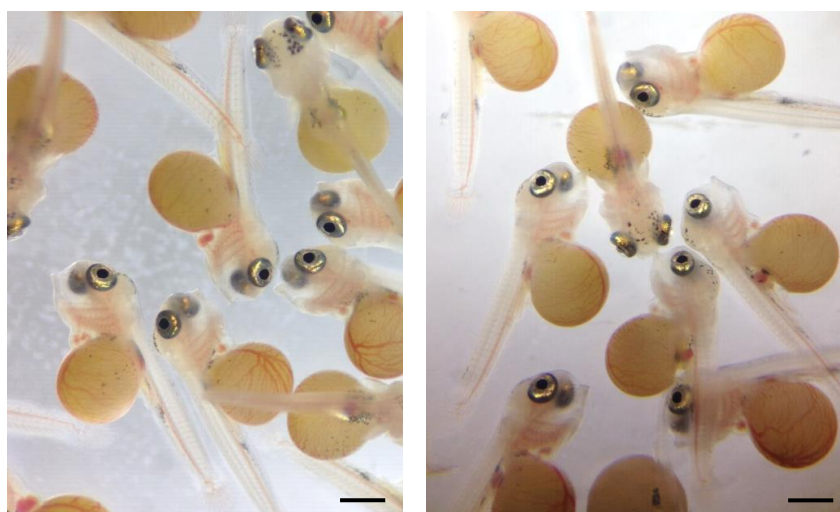
Figure 5.2. Photograph of live tilapia larvae at 3 dah. (A) Uninjected control larvae. (B) *piwil2* sgRNA/Cas9 injected larvae. Scale bar = 1 mm.

5.3.2 Impact of *piwil2* KO on PGCs survival (H&E)

The 52 confirmed *piwil2* mutant larvae by qPCR melt curve analysis were subjected to histological observation of PGCs using the serial transverse section of body cavity stained with H&E. Three different phenotypes were observed in *piwil2* mutants (Fig. 5.3, Table 5.3). In type A, no putative PGCs were observed (15/52, 29%, Table 5.3, Fig. 5.3A&B) and in type B, putative PGCs were morphologically atrophic and/or locally restricted (13/52, 25%, Table 5.3, Fig. 5.3C&D), while type C showed putative PGCs similar to the control (24/52, 46%, Table 5.3, Fig. 5.3E&F). More than half of "gene edited" larvae (type A&B, 28/52, 54%) showed a physiological impact of *piwil2* KO on PGCs.

Table 5.3. Phenotypes of gonadal anlagen of *piwil2* mutants induced by CRISPR/Cas9

Phenotype	Description	%
A	No gonadal anlagen and PGCs	15/52, 29%
B	Morphologically atrophic and/or locally restricted PGCs	13/52, 25%
C	Gonadal anlagen and PGCs similar to control	24/52, 46%
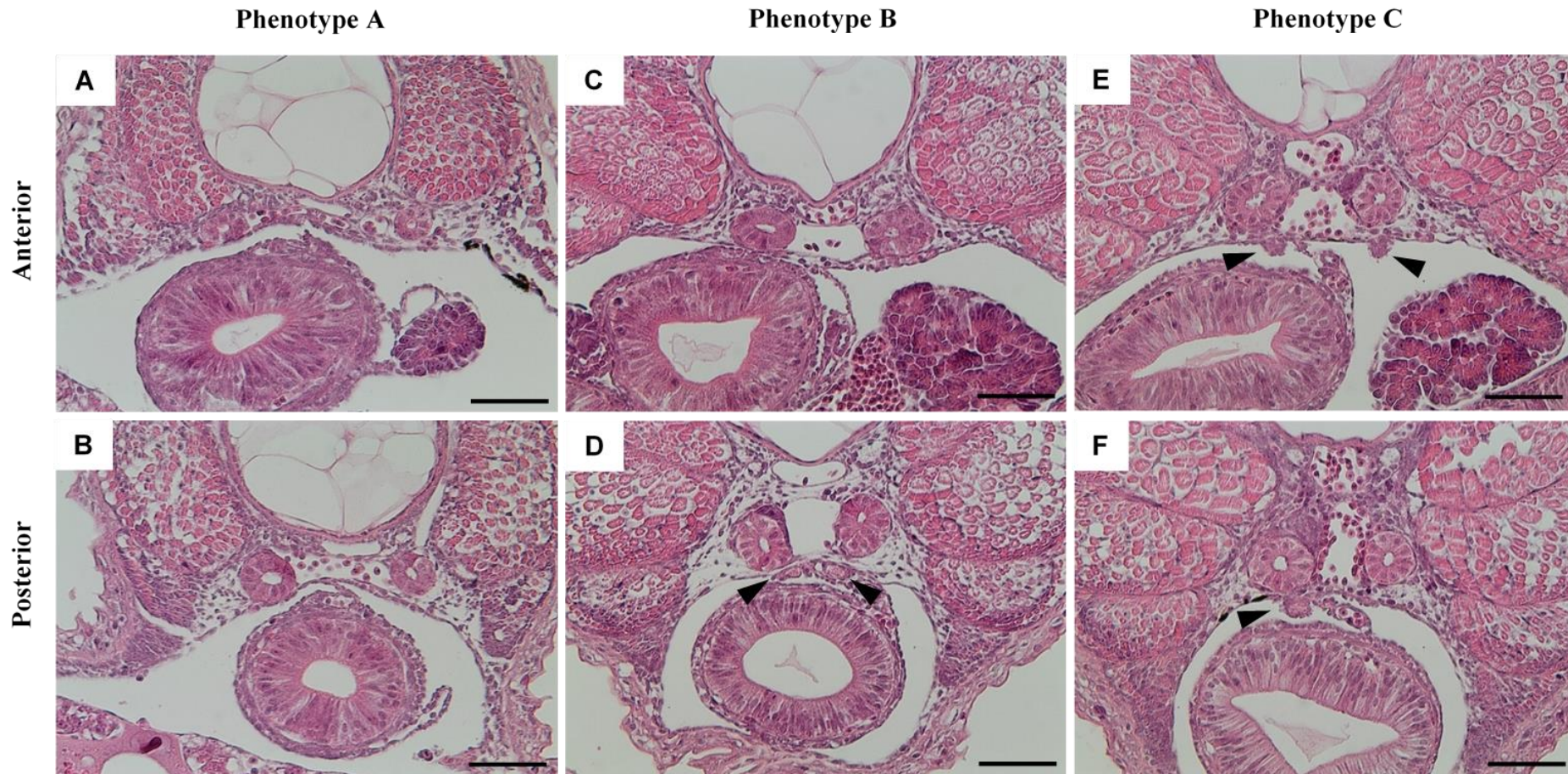
Figure 5.3. Histological observation of gonadal anlagen and PGCs in serial transverse sections of 3 dah *piwil2* mutants induced by CRISPR/Cas9. Three different phenotypes are shown: type A, no gonadal anlagen and PGCs (A&B); type B, morphologically atrophic and/or locally restricted PGCs (C&D); type C, gonadal anlagen and PGCs similar to control (E&F). Arrowheads indicate gonadal anlagen with or without PGCs. Scale bar = 50 μm.

5.3.3 ISH and expression level of *vasa*

vasa expression was detected by ISH in mature testis of Nile tilapia (Fig. 5.4A), with no signal detected in the negative controls hybridised with sense probe (Fig. 5.4B). However, in 3 dah larvae, PGCs were not visualised by ISH with antisense *vasa* riboprobes (data not shown). The relative number of *vasa* transcripts with respect to the number of *β-actin* transcripts analysed by absolute qPCR was 0.0282 ± 0.0016 in the testis ($n = 2$), while in 3 dah control larvae it was significantly lower ($p < 0.0001$), being three orders of magnitude lower at $2.09\mathrm{e}\text{-}05 \pm 6.55\mathrm{e}\text{-}06$ ($n = 4$ larvae).
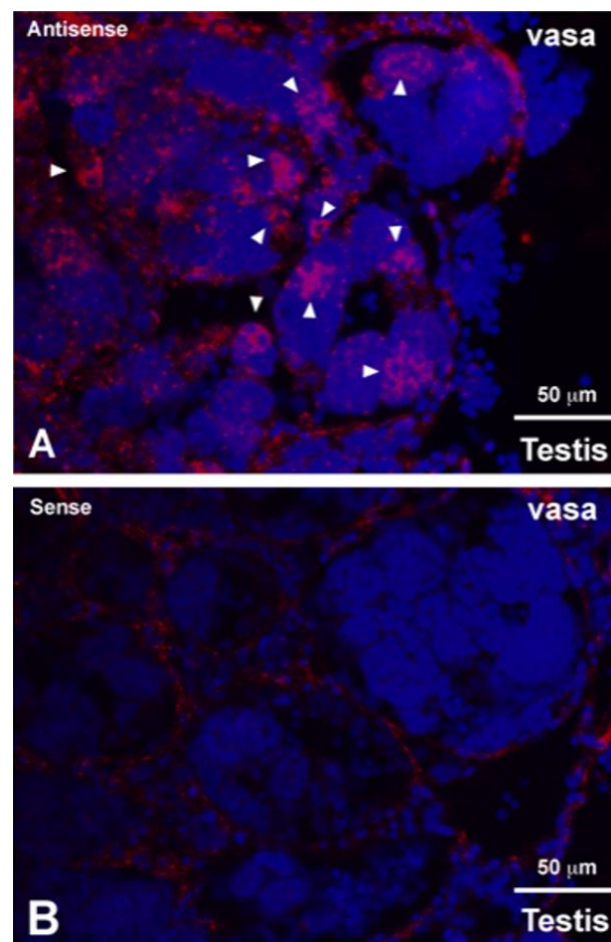
Figure 5.4. *vasa* expression in mature testis by fluorescent *in situ* hybridisation. Intense *vasa* expression was detected in the testis with antisense probe (A) and indicated with arrowheads, while lacking signal shown in the testis with sense probe (B).

5.3.4 Melt curves associated with phenotypes

Phenotype A *piwil2* mutants consisted of 20% and 80% of melt curve type I and II, respectively, and 0% of melt curve type III (Fig. 5.5). Phenotype B mutants showed 7.7%, 76.9% and 15.4% of melt curve type I, II and III, respectively (Fig. 5.5). In phenotype C, there were 91.7% and 8.3% of melt curve type II and III, respectively, but no melt curve type I (Fig. 5.5).
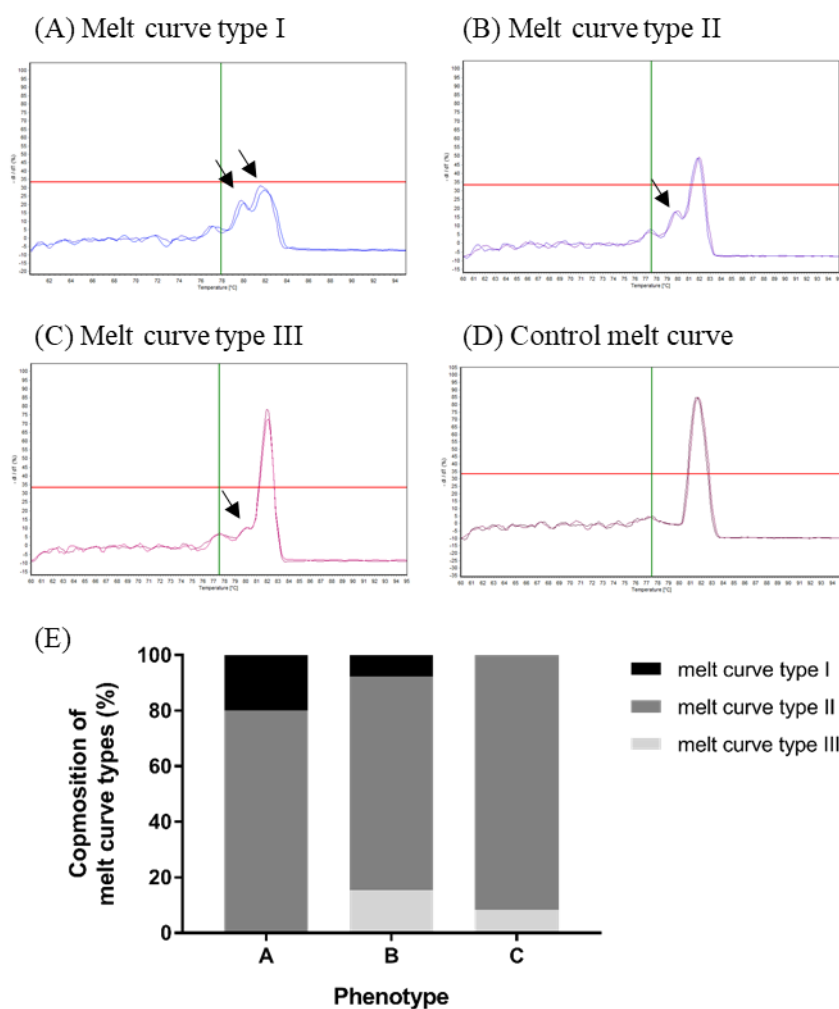
Figure 5.5. Composition of melt curve types in each phenotype of *piwil2* mutants. (A-D) Representative shapes of melt curves of each type: (A) Melt curve type I, two major peaks (arrow); (B) type II, one minor (arrow) and major peak; (C) type III, a small ramp (arrow) before the major peak; and (D) control melt curve, a single major peak. (E) Composition of melt curve types in each phenotype. Data are presented as percentage (%) of melt curve types in each phenotype.

5.4 Discussion

This Chapter aimed to test the functional effects in *piwil2* mutant produced by CRISPR/Cas9 genome editing method optimised in Chapter 4. The results showed gene editing resulting in >95% mutation rates. In addition, three different phenotypes were observed, based on PGC screening by histology, in *piwil2* KO tilapia larvae with more than half of the mutants showing either a complete absence of PGCs suggesting a putative sterile phenotype or PGCs morphologically atrophic and/or locally restricted PGCs.

In the current study, histological analysis revealed a different degree of phenotypic impacts on PGCs in *piwil2* KO mutant larvae. Three different phenotypes were found amongst mutants including phenotype A in which no PGCs could be detected by histology, phenotype B in which morphologically and/or locally restricted PGCs were observed; and phenotype C which was comparable to control individuals with PGCs apparently normal. The loss of germ cells and/or deficiency in germ cell development have been previously reported in *piwi* mutants in several species. In *Drosophila*, the loss of *piwi* resulted in a failure of the proliferation of germline stem cells (GSCs) (Lin and Spradling, 1997) suggesting a crucial role of *piwi* in regulating the number and division of GSCs (Cox *et al.*, 2000). In mice, three murine *piwi* family genes have been reported (*Miwi*, *Mili* and *Miwi2)* and *Miwi* null mutant resulted in the arrestment of spermatogenesis at the spermatid stage (Deng and Lin, 2002), while *Mili* and *Miwi2* null mutants showed a blockage of the first meiosis resulting in sterility (Kuramochi-Miyagawa *et al.*, 2004; Carmell *et al.*, 2007). In zebrafish, the loss of *ziwi* showed consecutive loss of germ cells due to apoptosis during early development, suggesting a role in the maintenance of germ cells (Houwing *et al.*, 2007). In addition, loss of *zili* also caused the loss of germ cells and the functions of *zili* in transposon defence, germ cell differentiation and meiosis were suggested in zebrafish (Houwing *et al.*, 2008). In medaka, *piwil1* and *piwil2* were also maternally deposited and expressed in PGCs in larvae, oogonia to mature oocytes in the ovary and spermatogonia to spermatocytes in the testes of adults (Zhao *et al.*, 2012). Knockdown of zygotic *piwil1* in medaka significantly reduced the number of PGCs and hindered the migration of PGCs but did not prevent the formation of PGCs (Li *et al.*, 2012). Together with the previous results on *piwi* loss of function studies, this suggests that the putative loss of PGCs observed in *piwil2* KO larvae (phenotype A) supports the reported role of the PIWI domain of *piwil2* gene in the maintenance of PGCs. However, while early assessment of phenotypic effects

in gene edited individuals can demonstrate a physiological role of the target gene, it does not mean the phenotype will result in sterile animals later during their life cycle. Therefore, further studies are required to assess the longer-term functional effect of the gene editing method later during the life cycle especially during puberty.

The impact of *piwil2* KO on PGCs was investigated by histological observation of gonadal anlagen at the early larval stage at 3 dah, which were located in the dorsal peritoneal wall. In Nile tilapia, on the day of hatching (0 dah), PGCs can be found at the outer layer of the lateral plate mesoderm around the hindgut, which are morphologically distinguishable from somatic cells (Ijiri *et al*., 2008; Kobayashi *et al*., 2000). Then, at 3 dah, PGCs are found in the gonadal anlagen after the formation of the coelomic cavity in the lateral plate mesoderm in Nile tilapia, and soon after PGCs starts to proliferate (Farlora *et al*., 2014; Ijiri *et al*., 2008; Kobayashi *et al*., 2000). However, results based on histological assessment must be interpreted with caution. Although phenotype A mutants appeared to not show PGC-like cells in any of the sections analysed, while all control fish showed PGC-like cells in the targeted sections, the possibility that some PGCs were missed in the histological screening of phenotype A mutants cannot be completely ruled out due to the serial sectioning strategy. Further assessment is, therefore, required to confirm phenotypes and conclude on the functional effect of the gene editing methodology.

Another method to investigate the effects of *piwil2* KO on PGCs would be the investigation of expression levels of germ cell marker genes. However, in the current study, it was not possible to conduct both histological assessment and gene expression analysis in the same individuals. To further confirm the presence and absence of PGCs in 3 dah larvae, ISH technique using *vasa* riboprobe was performed in this study. However, while gene expression was clearly localised in mature testis using the antisense *vasa* riboprobe, no signal could be detected in 3 dah larvae. Further absolute qPCR analysis revealed that the abundance of *vasa* transcripts in 3 dah larvae was significantly lower than in testis by three orders of magnitude. ISH, therefore, failed to detect the *vasa* signal on the histological section of 3 dah larvae due to the low abundance of *vasa* as well as the low sensitivity of the method. In previous studies, PGCs have been identified by ISH using *vasa* riboprobe (Kobayashi *et al*., 2000) and immunohistochemistry (IHC) using anti-Vasa antibody (Kobayashi *et al*., 2002) in Nile tilapia at the larval stage. Despite many attempts, we could not have access to the anti-Vasa antibody; therefore, could not pursue the IHC route. However, PGCs could be observed histologically with

H&E staining in this study as previously reported in various fish species including zebrafish (Hashimoto *et al*., 2004), medaka (Sawamura *et al*., 2017), goldfish (Goto *et al*., 2012), grass puffer (*Takifugu niphobles*) (Hamasaki *et al*., 2017), brown-marbled grouper (*Epinephelus fuscoguttatus*) (Boonanuntanasarn *et al*., 2016), yellowtail kingfish (*Seriola lalandi*) (Fernández *et al*., 2015) and Nile tilapia (Farlora *et al*., 2014; Ijiri *et al*., 2008). PGCs can be distinguished by their morphological shape with a large and round nucleus, a weak staining by haematoxylin and their location.

A range of mutations was apparent in the *piwil2* edited individuals screened by qPCR melt curve analysis, but further genotyping analysis is required to better characterise the genotypes and correlate these to the observed phenotypes. The qPCR melt curve analysis method relies on the presence of the heteroduplex formed between WT and mutated sequence or two differently mutated sequences which cause the changes in the melt temperature and the corresponding melt curve (Chateigner-Boutin and Small, 2007). Mutants with the greatest deviation in their melt curve (melt curve type I) were more prevalent in phenotypes A and B but were not observed in phenotype C grouping. This method, however, relies on a subjective decision as the melt curves are a continuum and it was unable to predict accurately the observed phenotypic difference due to its limited ability to inform on the individuals complex genotype. Further genotyping assessment is required such as deep sequencing to characterise the relationship between the genotypes and the different degree of phenotypes observed in this study (Zischewski *et al*., 2017).

In the current study, *piwil2* KO mutants developed normally without significant difference in survival rate at early larval stages compared to the controls which indicates that *piwil2* gene is not essential for normal larval development. In mice, it has equally been shown that embryonic somatic cell division is normal in *Mili* null mutants (Kuramochi-Miyagawa *et al*., 2004). However, further assessment is required to determine potential pleiotropic effects in Nile tilapia over a longer time frame as it was suggested in Atlantic cod which showed multiple transcriptional changes by knockdown of one of the GP components, *dnd* (Škugor *et al*., 2014b).

In conclusions, this study showed putative germ cell loss was induced by KO Piwi domain in *piwil2* using CRISPR/Cas9 system in Nile tilapia. A high mutation rate (*circa* 96%) was achieved in *piwil2* mutants without significant treatment mortality demonstrating the reproducibility of the optimised protocol developed in Chapter 4. Histological observations revealed that more than half of injected tilapia eggs developed

into *piwil2* KO larvae with a phenotypic effect on PGCs. This result implies that the Piwi domain of *piwil2* plays an important role in the survival of PGCs during early development of Nile tilapia larvae. However, it is still very early days in the development of this potential novel stock management strategy and further investigations are required to confirm the long-term physiological effect in *piwil2* KO fish (*e.g.* sterility), characterise the range of resulting genotypes and the association with the different phenotypes found in *piwil2* mutants.

CHAPTER 6

COMPARISON OF METHODS TO STUDY THE GENOTYPE OF MOSAIC *PIWIL2* MUTANT NILE TILAPIA INDUCED BY CRISPR/CAS9

6.1 Introduction

It was evident from the previous Chapter 5 that the physiological impact of genome editing is not necessarily reproducible and clearly does not associate with a simple classification of "mutant" or "not", as assayed by simple qPCR melt curve analysis. It is, therefore, necessary to undertake a more in-depth consideration of the genome editing effect and subsequent individual genotype diversity present following the CRISPR/Cas9 editing.

CRISPR/Cas9 is an adaptive immune system against viral infection found in bacteria and archaea (Lander, 2016; Makarova *et al*., 2015). Cas9 belongs to Class 2, type II Cas effector, which is a single protein effector derived from either *S. pyogenes* or *S. thermophilus* and it has two nuclease motifs, RuvC and HNH, producing DSBs at 3 nucleotides upstream of the PAM (Deveau *et al*., 2008; Makarova *et al*., 2006). The cut sites are then repaired by either an error-prone NHEJ or HDR. While NHEJ introduces indel into the vicinity of the cleaved site, HDR can edit the cut site precisely by adding a donor DNA which has a desirable sequence (Huang *et al*., 2011; Rodgers and McVey, 2016), leading to a precise genome editing, including targeted gene insertion, correction and point mutagenesis (Kim and Kim, 2014). However, NHEJ repair pathway is more dominant than HDR, as HDR occurs only in S and G2 phase while NHEJ occurs in the entire cell cycle (Deriano and Roth, 2013). The efficiency of HDR needs to be improved for applications such as functional gene analysis in injected animals (F0). Even though the high efficiency of CRISPR/Cas9 mediated genome editing can produce biallelic mutants in founder (F0) individuals (Jao *et al*., 2013), it is widely recognised that there is a high mosaicism present in F0 animals created by CRISPR/Cas9 and NHEJ DNA repair. This is an emerging challenge which needs to be studied for the direct phenotypic analysis in F0 (Trubiroha *et al*., 2018).

Incomplete genome editing due to the low translation activity for Cas9 protein in the zygote, the deferred function of sgRNA/Cas9 complex and the subsequent DNA repair beyond the one-cell stage embryo, can lead to mosaicism which is a known phenomenon resulting in cells within the same individual having different genotypes (Hsu *et al*., 2014; Oliver *et al*., 2015; Yen *et al*., 2014). The uncontrollable impact of mosaicism may prevent direct phenotypic analysis and interpretation of gene function in F0 animals. However, the lack of comprehensive and efficient genotype screening methods for the F0 animals is a technical hurdle that must be overcome in mutagenesis

studies using CRISPR/Cas9. To this end, reliable and efficient genotype screening methods for F0 animals need to be validated to properly screen for complex genotypes in treated individuals.

In the previous Chapter 5, melt curve analysis of PCR amplified region spanning the CRISPR/Cas9 attack site on the *piwil2* gene was performed using routine quantitative PCR. This has the advantage of being a relatively simple method, rapid and cost effective, but it cannot provide further details on the potential complex genotypes which could drive the differential phenotypic response observed. This method, however, is widely used in the literature to screen on-target mutation in F0 animals (Dahlem *et al*., 2012; Zhang *et al*., 2015b). The most widely used assay is mismatch cleavage assay, which utilizes SURVEYOR endonuclease or T7E1 to cleave the mismatched extrahelical loops generated by reannealing between WT and mutated target fragments or between differently edited mutated fragments with treated samples being visualised by gel electrophoresis (Vouillot *et al*., 2015). It should be noted, however, that SURVEYOR or T7E1 cannot detect a biallelic homozygous mutation that can result in a false negative result and the presence of polymorphisms can cause a false positive result (Fig. 6.1) (Kim *et al*., 2014a). CRISPR/Cas-derived RNA-guided engineered nucleases (RGENs) uses a similar analytical approach, but unlike T7E1 it can distinguish monoallelic, heterozygous and homozygous biallelic mutations by digesting the unmodified target site (Fig. 6.1) and this is not affected by the polymorphisms adjacent to the target site (Kim *et al*., 2014a). Restriction enzyme digestion (RED) also digests the unmodified target sequence in the PCR amplicon and visualised by gel, but this method is only available when the target site has the specific restriction enzyme site (Zischewski *et al*., 2017). Melt curve analysis as employed in Chapters 4 & 5, or the more advanced high resolution melt curve analysis (HRMA), are high-throughput screening methods which detect mutation by analysing the melt temperature shifts and melt curve shapes using qPCR apparatus (Thomas *et al*., 2014). HRMA differs from normal melt curve analysis by measuring at an order of magnitude greater resolution (*circa* 0.04°C vs. 0.5°C) on specific analysis platforms using proprietary chemistries, but this does allow the detection of single base mutations in amplified fragments (D'Agostino *et al*., 2016; Dahlem *et al*., 2012; Denbow *et al*., 2017). Fluorescent PCR-capillary gel electrophoresis (fragment analysis) is also a high-throughput method, which uses the differences in mobility of DNA fragments in capillary gel electrophoresis to characterise the abundance of fragment sizes in a given sample and like HRMA, is capable of differentiating single-nucleotide difference (Ramlee *et al*.,

2015). Among all the genotyping methods, Next Generation Sequencing (NGS) of the target region is the most comprehensive analysis method, providing the direct and detailed information about the nature and diversity of mutations with high sensitivity and high throughput format (Zischewski *et al*., 2017); however, it does require access to specialist sequencing platforms that are costly. Each of these techniques has drawbacks that hamper their use in a high-throughput format (Chapter 1, Table 1.2). Some of these assays are expensive, time consuming, lack sequence specificity, require costly capital equipment, or fail to distinguish biallelic combinations (Chapter 1, Table 1.2).

In sterility studies utilising genome editing techniques, there is a difficulty to produce homozygous progeny due to the induced sterility in F0. Therefore, it is essential to understand the complicated mosaic genotypes and correlate these to phenotypic effects. However, it was noted that so far, the genotyping methods for F0 fish in sterility studies were biased without comprehensive validation of the methodology. In CRISPR/Cas9-mediated sterility studies, *nanos2*, *nanos3*, and *eEF1A1b* KO sterile tilapia were screened by RED and Sanger sequencing in a very limited scale (Chen *et al*., 2017; Li *et al*., 2014), *dnd* and microRNA-202 KO sterile medaka by Sanger sequencing in a small scale (Gay *et al*., 2018; Sawamura *et al*., 2017) and *dnd* KO sterile salmon by PCR and Sanger sequencing of PCR products (Wargelius *et al*., 2016). In addition, TALEN-mediated *fshr* and *lhcgr* KO sterile zebrafish were screened by HRMA while ZFN-mediated *lhb* KO sterile Channel catfish were screened by SURVEYOR and Sanger sequencing (Qin *et al*., 2016; Zhang *et al*., 2015b). In particular, genome editing studies conducted in Nile tilapia targeting gonad-related genes has been limited to RED and Sanger sequencing (Feng *et al*., 2015; Jiang *et al*., 2016, 2017; Li *et al*., 2015; Xie *et al*., 2016; Zhang *et al*., 2014). There is a clear gap in the understanding of the genotypes in mutant fish founders due to the lack of comprehensive validation of the screening methodologies.

The outputs of Chapter 5 provided a timely opportunity to compare in the current Chapter the most frequently used mutation screening methods, T7E1, RGEN, HRMA, fragment analysis and NGS, with the aim to validate an effective and efficient high-throughput screening method. If future research is to continue to apply CRISPR/Cas9 and examine phenotypic response in F0 animals, it is necessary to operate at the individual level to be able to accurately characterise the individual's potentially complex genotype and relate this to, and ideally even predict, the phenotypic response. The outputs of Chapter 5 allow such a comparison of methods to be performed and furthermore, the outputs from such a study (in particular NGS outputs) have the potential to improve our

understanding of the link between the genotypic differences associated with the observed phenotypes in CRISPR/Cas9 *piwil2* mutants.

6.2 Materials and Methods

6.2.1 *piwil2* CRISPR/Cas9 mutant larvae

gDNAs were extracted from the head and tail of eight WT controls and fifty-two *piwil2* KO Nile tilapia larvae (previously generated in Chapter 5) using SSTNE and stored at -20˚C. In Chapter 5, the genotype (mutant or not) of each individual had been confirmed by regular qPCR melt curve analysis and following histological examination of the gastro-intestinal trunk to screen for the presence and abundance of PGCs. Fifty-two mutants were classified into three phenotypes: (1) phenotype A ($n = 15$, mutant sample No. 1 to 15), no apparent gonadal anlagen and PGCs; (2) phenotype B ($n = 13$, mutant sample No. 16 to 28), PGCs appeared morphologically atrophic and/or locally restricted; (3) phenotype C ($n = 24$, mutant sample No. 29 to 52), gonadal anlagen and PGCs similar to control.

(A)

(B)

(C)
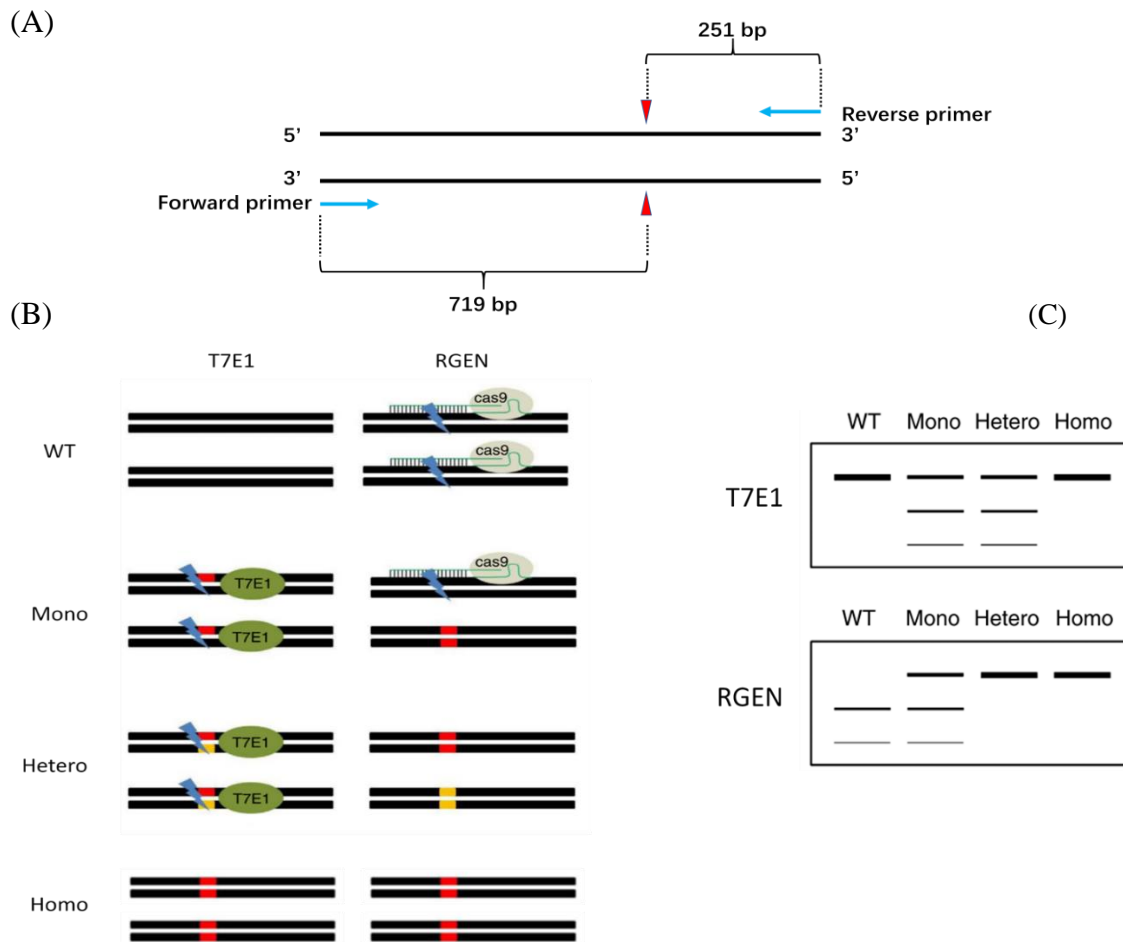
Figure 6.1. Overview of T7E1 and RGEN assays. (A) Primer design for T7E1 and RGEN assays. The primer pairs were designed asymmetrically flanking to the canonical cut site of CRISPR/Cas9 (arrowheads) with 719 bp on the left and 251 bp on the right. (B) Schematic overview of the different cleavage results between T7E1 and RGEN assay in four possible scenarios in a diploid cell: (WT) wild type, (Mono) a monoallelic mutation, (Hetero) heterozygous biallelic mutations and (Homo) homozygous biallelic mutations. Black lines represent PCR products derived from each allele; yellow and red boxes indicate indel mutations generated by NHEJ. (C) Expected results of T7E1 and RGEN digestion resolved by electrophoresis. RGEN can distinguish monoallelic, heterozygous and homozygous biallelic mutations, while T7E1 cannot detect the biallelic homozygous mutation. The conceptual diagrams of B and C were obtained from Kim *et al*. (2014a).

6.2.2 Mutation screening methods

In this study, five different methodologies were used to screen for mutations in *piwil2* CRISPR/Cas9 mutants larvae including T7EI, RGEN, HRMA, fragment analysis and NGS, with the aim to validate an effective and efficient high-throughput screening method.

6.2.2.1 T7EI and RGEN assays

- PCR and purification

The primer pair, *piwil2* sgRNA2_2, was designed using NCBI primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) to amplify a product which has the *piwil2* sgRNA2 target site offset from the centre of the amplicon so that the digested fragments are easily resolvable on the gel electrophoresis (Fig. 6.1, Table 6.1). No polymorphic variations in the amplicon sequence were found in NCBI. The PCR reaction run in a total volume of 40 µL, consisted of 20 µL of Q5 Hot start High-fidelity 2X master mix (NEB, Ipswich, USA), 2 µL of 10 µM of forward and reverse primer, respectively, 80 ng of gDNA and Milli-Q water up to 40 µL. The PCR program was: 98˚C for 30 sec, followed by 35 cycles of 98˚C for 5 sec, 67.5˚C for 10 sec and 72˚C for 20 sec, with a final extension at 72˚C for 2 min, and conducted by TGradient thermal cycler (Biometra GmbH, Göttingen, Germany). Following the PCR, each PCR product was purified using GeneJET PCR purification kit (Thermo Fisher, Waltham, USA) and the amplicon DNA size was confirmed by electrophoresis on a 1% agarose gel stained with EtBr and quantified using nanodrop. The size of amplified fragments (both target sequence as well as unexpected putative large indel fragments) was calculated in comparison to a DNA size standard ladder (GeneRuler 1 kb DNA Ladder, Thermo Fisher, Waltham, USA) from the digital image of the gel electrophoresis using GeneTools (Syngene, Cambridge, UK).

Table 6.1. Primer list used for mutant screening

Primer pair		Sequence of primer (5'-3')	Amplicon size of WT	Annealing temp (°C)	Purpose
piwil2 sgRNA2_2	F	ACCTGTGCCGTAAGGCTGGA	970 bp	67.5	T7E1 assay, RGEN assay
	R	AGTGTGCAGAAAACACTGACTTCAC			
piwil2 sgRNA2 HRM	F	TAGGTGAGAATTAGGTGTGGTTT	211 bp	62-57 (touchdown)	HRMA
	R	TGCACAATGCATGAGTCCTAC			
piwil2 sgRNA2_ CAG	F	CAGTCGGGCGTCATCATAGGTGAGAATTAG GTGTGGTTT	227 bp	56.7	Fragment analysis
	R	TGCACAATGCATGAGTCCTAC			
piwil2 sgRNA2 NGS	F	TCGTCGGCAGCGTCAGATGTGTATAAGAGA CAGTAGGTGAGAATTAGGTGTGGTT*T	282 bp	55.5	NGS
	R	GTCTCGTGGGCTCGGAGATGTGTATAAGAG ACAGTGCACAATGCATGAGTCCTA*C			

Grey shade, CAG tailing sequence; boxed sequences, Illumina overhang adapter sequences. The primer pair for HRMA and NGS were purified by HPLC. NGS primer pair has 3' modification. Asterisk (*) denotes a phosphorothioate (PTO) bond.

- T7 endonuclease I (T7EI) assay

Prior to T7E1 digestion, 100 ng of each purified amplicon was hybridised to allow to form heteroduplex in 10 μL reaction volume containing 1 μL of NEBuffer 2 (NEB, Ipswich, USA) and MilliQ water up to 10 μL. The thermocycler condition for hybridisation was as follows: initial denaturation at 95˚C for 5 min, annealing from 95˚C to 85˚C at -2˚C/sec ramp rate and 85˚C to 25˚C at -0.1˚C /sec ramp rate and termination at 20˚C. The hybridised product was digested by 0.5 μL of T7EI (NEB, Ipswich, USA) at 37˚C for 15 min and terminated by adding 1 μL of 0.25 M EDTA. The fragmented PCR products were then run on the agarose gel and the percent of nuclease-specific cleavage products were determined by GeneTools (Syngene, Cambridge, UK). The cleavage efficiency of T7E1 was calculated by the following formula (Guschin *et al.*, 2010) and used as the estimated arbitrary gene modification rate:

- Arbitrary gene modification rate (%) = $100 \times (1 - (1 - \text{fraction cleaved})^{1/2})$
- Fraction cleaved = sum of cleaved band intensities/sum of cleaved and parental band intensities

The cleavage at the target site of *piwil2* sgRNA2 produces two cleaved fragments of 719 and 251 bp in size. The imaging condition of gel bands was optimised to obtain high sensitivity of the DNA fragments dyed with EtBr by adjusting the exposure time (40 ms) and the loading volume of the digested product (40 ng) based on the

measurement of the band intensity of the serial dilution of 1 kb and 100 bp DNA ladder (Thermo Fisher, Waltham, USA). The reproducibility of this method was tested by analysing nine randomly selected mutant samples in triplicate on the gel. The intra- and inter-assay CVs were 1.98 and 3.95%, respectively.

- CRISPR/ Cas-derived RNA-guided engineered nuclease (RGEN) assay

The *piwi2* sgRNA2 and Cas9 nuclease protein (*S. pyogenes*) (NEB, Ipswich, USA) were used to examine the mutation efficiency of *piwil2* sgRNA2 mutants, according to the manufacturer's protocol. 50 nM of *piwil2* sgRNA2 and 50 nM of Cas9 nuclease were incubated at 25˚C for 10 min with a provided reaction buffer (NEB, Ipswich, USA) to assemble the Cas9/sgRNA complex. Then, the purified amplicons were added at the final concentration of 4.5 nM as the substrate DNAs and incubated at 37˚C for 16 hours. The assay included eight positive controls of WT samples and three negative controls of no DNA substrates. To obtain the best cleavage efficiency, the molar ratio of Cas9 and sgRNA per target site was kept at 11:11:1 in a total reaction volume of 10 μL. The cleavage reaction was terminated by incubating at 80˚C for 5 min. The fragmented PCR products were then resolved with 1% agarose gel electrophoresis and visualised with EtBr staining. The percentage of nuclease-specific cleavage products were determined by measuring each band intensity using GeneTools (Syngene, Cambridge, UK). The estimated arbitrary gene modification rate analysed by RGEN digestion was calculated by the following formula modified from Kim *et al*. (2014a):

- Arbitrary gene modification rate (%) = 100 × (intensity of uncleaved band (*i.e.* parental band)/ (intensity of uncleaved band + intensity of large cleaved band + 0.3491 × intensity of large cleaved band)

The large and small cleaved fragments are 719 and 251 bp in size. The small cleaved fragments of 251 bp were masked by similar sized artefacts (200 – 300 bp) of the RGEN reaction. Such artefacts were also found in control reaction which has no Cas9 enzyme, but sgRNA, gDNA and the buffer, indicating that these are not cleaved DNA, but artefacts (data not shown). Thus, the small cleaved fraction was estimated based on the intensity of the corresponding large cleaved fraction by multiplying the proportional length which is 0.3491. This calculation assumes that the molar ratio between the cleaved large and small fragments are 1:1 and the band intensities are proportional to the length of the fragment as EtBr binds dsDNA once every 4-5 bp without apparent sequence preference (Waring, 1965). The imaging condition of gel bands was optimised to obtain

high sensitivity of the DNA fragments dyed with EtBr by adjusting the exposure time (80 ms) and the loading volume of the digested product (18.9 ng) by measuring the band intensity of the serial dilution of 1 kb and 100 bp DNA ladder (Thermo Fisher, Waltham, USA). The reproducibility of this method was tested by analysing nine randomly selected mutant samples in triplicate on the same gel. The intra- and inter-assay CVs were 1.96 and 3.50%, respectively.

6.2.2.2 High resolution melt curve assay (HRMA)

Primer pair for HRMA was designed using NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and HPLC purified (Eurofins, Brussels, Belgium) (Table 6.1). No polymorphic variations in the amplicon sequence were found in NCBI. The PCR reaction for HRMA contained 5 μL of 2X LightCycler 480 High Resolution Melting Master Mix (Roche, Basel, Switzerland), 1.2 μL of 25 mM MgCl2, 0.3 μL each 10 μM primer, 25 ng of gDNA and Milli-Q water up to 10 μL. Each mutant and control sample was tested in triplicate with four non-template controls (ntc) all contained within a single 384 well plate analysed using the LightCycler® 480 Instrument (Roche, Basel, Switzerland). The PCR program was: pre-incubation at 95℃ for 10 min, 45 cycles of denaturation at 95℃ for 15 sec, touchdown annealing (62℃ to 57℃ with 0.5℃ decrement/cycle) for 15 sec and extension at 72℃ for 15 sec. Then, it was followed by the HRMA program: 95℃ for 1 min, 40℃ for 1 min, 70℃ to 95℃ with 25 acquisitions per degree centigrade (acquisition at 0.04°C increments). The result was analysed by Gene Scanning and Tm calling analyses in LightCycler® 480 Software. The reproducibility of this assay was assessed by running the same HRMA of 52 mutants, 8 controls twice in triplicate. The intra- and inter-assay CVs of melt temperature (Tm) were 0.02 and 0.08%, respectively.

6.2.2.3 Fragment analysis

- Fluorescent labelled tailing PCR

To carry out fragment analysis of mutations in target site, PCR was performed using a fluorescent labelled tailed primer method (Boutin-Ganache *et al*., 2001). This method employs two different forward primers and one reverse primer. The initial PCR is to incorporate CAG sequence into the PCR product using CAG tailed forward primer and the reverse primer pair. The resultant tailed PCR products are used as a template for the

second stage PCR. A labelled primer with CAG dye replaces the tailed forward primer to amplify PCR product containing the fluorescent dye, which can be detected by the capillary sequencer (Beckman Coulter, Brea, USA). In this study, one type of dye, CAG_green (5'Dye-CAGTCGGGCGTCATCA-3') (Sigma-Aldrich, St. Louis, USA), was used for mutations created by *piwil2* sgRNA2. This dye sequence was added to the 5' prime end of the forward primer and paired with non-tailed reverse primer (Table 6.1). The PCR reactions run in a total volume of 8 µL, consisted of 4 µL of Q5 Hot start High-fidelity 2X master mix (NEB, Ipswich, USA), 0.15 µL of 1 µM tailed forward primer, 0.25 µL of 10 µM non-tailed reverse primer and 0.25 µL of 10 µM fluorescent dye labelled primer, 25 ng of gDNA and Milli-Q water up to 8 µL. PCR program was: 98˚C for 30 sec, followed by 33 cycles of 98˚C for 10 sec, 62˚C for 20 sec and 72˚C for 20 sec, with a final extension at 72˚C for 2 min, and conducted on a TGradient thermal cycler (Biometra GmbH, Göttingen, Germany). Following the fluorescent labelled tailing PCR, the PCR product size should be increased by 16 bp for CAG tailing, and the single band of expected increased product size compared with control samples was confirmed by electrophoresis on a 1% agarose gel stained with EtBr.

- Size determination of fragments

Size determination of the fluorescently labelled PCR products was assessed using a Beckman Coulter CEQ8000 sequencer (Beckman Coulter, Brea, USA) and associated software. For each capillary run, 0.7 - 1.5 µL of the tailed PCR product was added into an opaque 96 well sequence plate with V bottom (Beckman Coulter, Brea, USA) containing 30 µL of SLS and 0.4 µL of DNA Size Standard kit-400 (Beckman Coulter, Brea, USA). The added volume of each tailed PCR products was decided based on the intensity of the band on the gel electrophoresis. One drop of mineral oil was added to each well to avoid evaporation and the plate was briefly spun down. A 96-well electrophoresis buffer plate with a flat bottom (Beckman Coulter, Brea, USA) was prepared and contained 300 µL of A.C.E. Sequencing buffer (VWR Chemicals, Radnor, USA) in the corresponding wells. To analyse samples, Beckman Frag-3 (size range 60-400bp, Beckman Coulter, Brea, USA), 5 min pause program was used. Once the run was completed the data was analysed via the Fragment Analysis Module. All the obtained fragment lengths from the module were standardised based on the WT fragment length in controls of 227.69 ± 0.04 nt ($n = 8$) with indel size thereafter being described as + or − WT length with the indel size values being rounded off to the nearest whole nucleotide number. The proportion of each fragment within the mosaic genotype was calculated

based on the height of the fragment (Carrington *et al*., 2015). The estimated arbitrary gene modification rate assessed by fragment analysis was calculated by the following formula:

- Arbitrary gene modification rate (%) = 100 – proportional height of zero indel fragment (%)

6.2.2.4 Next generation sequencing (NGS)

- Sequencing library preparation and sequencing

Sequencing libraries of the target region of *piwil2* sgRNA2 from both control and putative mutants gDNA samples were prepared according to the Illumina MiSeq system instructions (Fig. 6.2). In the first PCR, the target region of *piwil2* gene was amplified with *piwil2* sgRNA2 NGS primer which has overhang adapters in each 5' prime to incorporate the overhang adapters (Table 6.1). The first PCR reaction consisted of 10 µL of Q5 Hot start High-fidelity 2X master mix (NEB, Ipswich, USA), 0.3 µM each forward and reverse primer, 1.0 µL of gDNA (*circa* 100 ng) and MilliQ water up to 20 µL. The PCR thermal cycle was: initial denaturation at 98°C for 30 sec, 20 cycles of 98°C 10 sec, 55.5°C for 30 sec and 72°C for 30 sec and final extension at 72°C for 2 min. The amplicons were purified using AxyPrep Mag^{TM} PCR clean-up kit (Axygen, Corning, USA) and eluted in 12 µL of MilliQ water. In the second PCR, two indices and Illumina sequencing adapters were added to each amplicon using Nextera XT index primers. The index PCR reaction comprising 12.5 µL of Q5 Hot start High-fidelity 2X master mix (NEB, Ipswich, USA), 2.5 µL of each Nextera XT Index primer, 2.5 µL of the purified PCR product and MilliQ water up to 25 µL. The thermal cycle of index PCR was: initial denaturation at 95°C for 30 sec, 10 cycles of 95°C 10 sec, 55°C for 30 sec and 72°C for 30 sec and final extension at 72°C for 5 min. 1 µL of each PCR product was run on the agarose gel stained with EtBr to check the size and band intensity and 1 more amplification cycle was added, if necessary (mutant No.11 and control No. 6). The amplicons were purified using AxyPrep Mag^{TM} PCR clean-up kit (Axygen, Corning, USA) and eluted in 10 µL of MilliQ water. Each sample was quantified by using the Qubit® 2.0 Fluorometer and dsDNA HS assay kit (Thermo Fisher, Waltham, USA) and added into the library at the final concentration of 1 nM. The library was quantified by using the Qubit® 2.0 Fluorometer and dsDNA HS assay kit (Thermo Fisher, Waltham, USA) and normalised to the final concentration of 10 nM. Given the relatively low

amount of coverage required per individual, it was considered uneconomical to sequence the library in a dedicated sequencing run. Therefore, the double-indexed library was combined with an existing MiSeq sequencing run being undertaken for fish egg microbiome identification. Libraries were sequenced on an Illumina MiSeq instrument using a MiSeq reagent kit v2 of 250 bp paired-end reads with the sequencing run comprising of 91.5% 16S egg library, 2.5% CRISPR library and 6% phiX library (control). In total 20.5 M paired-end reads (Table 6.2) were produced of which 0.66 M paired-end reads belonged to the CRISPR study with the average number of reads per sample being $10,943 \pm 203$ per sample.

Table 6.2. Sequencing quality metrics summary

	Read 1 (Forward)	Read 2 (Index)	Read 3 (Index)	Read 4 (Reverse)
Lane	1	1	1	1
Tiles	28	28	28	28
Density (K/mm2)	987 ± 58	987 ± 58	987 ± 58	987 ± 58
Cluster PF (%)	84.92 ± 2.04	84.92 ± 2.04	84.92 ± 2.04	84.92 ± 2.04
Phas/Prephas (%)	0.033 / 0.014	0.000 / 0.000	0.000 / 0.000	0.167 / 0.345
Reads (M)	20.53	20.53	20.53	20.53
Reads PF (M)	17.42	17.42	17.42	17.42
% >= Q30	86.13	81.65	79.66	85.15
Yield (G)	4.35	0.12	0.12	4.35
Cycles Err Rated	250	0	0	250
Aligned (%)	4.12 ± 0.18	0.00 ± 0.00	0.00 ± 0.00	4.11 ± 0.18
Error Rate (%)	0.72 ± 0.06	0.00 ± 0.00	0.00 ± 0.00	0.88 ± 0.08
Error Rate, 35 cycle (%)	0.38 ± 0.07	0.00 ± 0.00	0.00 ± 0.00	0.26 ± 0.38
Error Rate, 75 cycle (%)	0.33 ± 0.19	0.00 ± 0.00	0.00 ± 0.00	0.22 ± 0.18
Error Rate, 100 cycle (%)	0.32 ± 0.15	0.00 ± 0.00	0.00 ± 0.00	0.23 ± 0.13
Intensity Cycle 1	32 ± 3	245 ± 33	199 ± 27	27 ± 3

Tiles—The number of tiles per lane.

Density—The number of clusters per mm2 detected by image analysis.

Clusters PF—The percentage of clusters passing filtering.

Phas./Prephas.—The value used by Real-Time Analysis (RTA) for the percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read.

Yield—The number of bases sequenced, in gigabases.

Cycles Err Rated—The number of cycles that have been error rated starting at cycle 1.

Aligned—The percentage of the sample that aligned to the PhiX genome.

Error Rate—The calculated error rate, as determined by the PhiX alignment. Subsequent number of cycles display the error rate for cycles 1–35, 1–75, and 1–100.

Intensity Cycle 1—The average of the four intensities (one per channel or base type) measured at the first cycle averaged over filtered cluster.
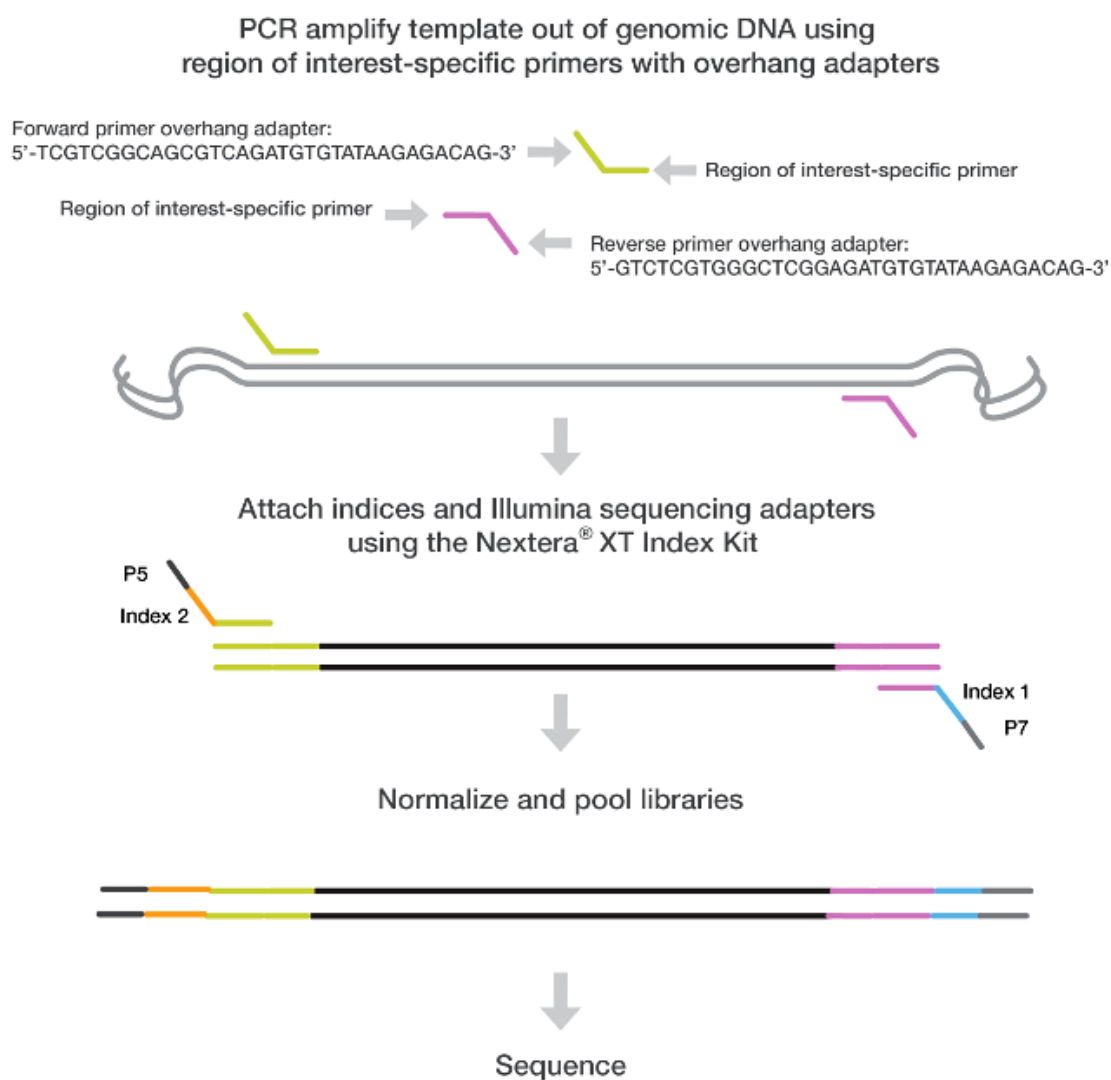
Figure 6.2. Workflow for preparation of amplicon for NGS. Overhang adapters are added to gene-specific primers and are used to amplify templates incorporating the overhang adapters from gDNA. This is used as template DNA in a subsequent limited-cycle amplification PCR to add multiplexing indices and Illumina sequencing adapters. Libraries are normalised and pooled and sequenced on the MiSeq system. This figure is originated from the Illumina MiSeq system instruction.

- Analysis of sequencing data using CRISPResso

FASTQ files generated by Illumina sequencing were analysed with CRISPResso, an analysis tool of CRISPR/Cas9 genome editing outcomes from deep sequencing data (Pinello *et al*., 2016). CRISPResso automates the following steps: 1. filtering low quality reads (phred33), 2. trimming adapters using Trimmomatic (Bolger *et al*., 2014), 3. merging paired-end sequences with FLASh (http://ccb.jhu.edu/software/FLASH/), 4. aligning the reads to a reference amplicon with needle from the EMBOSS suite, which is an optimal global sequence aligner based on the Needleman-Wunsch algorithm (Rice *et al*., 2000), 5. quantifying the proportion of NHEJ outcomes and 6. determining the proportion of frame-shift and in-frame mutations as well as detecting potential splice site mutations (Pinello *et al*., 2016). This pipeline was used for assessment of on-target editing efficacy, the percentage of WT, frequency of mutated alleles among NGS reads, spectrum of indels created by CRISPR/Cas9 editing events.

Paired FASTQ files were analysed using CRISPResso, applying the following parameters:

(1) WT sequence:

TAGGTGAGAATTAGGTGTGGTTTACAGAATCACAACACATTTCTGAAATGCATT
TGTACCTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGG
ATCAGTTCCTCATTGGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAAT
GAGCTGGAAGGGGGACAACACACTGCTGGTAGGACTCATGCATTGTGCA,

(2) sgRNA sequence, ACGGATCAGTTCCTCATTGG,

(3) coding sequence,

CTCATAGTACTTCTGCAGTGCAGCCAGTAAGCAAACCCTGAAGCCACGGATCAG
TTCCTCATTGGGGGTCTGGAAAATTACTCTGGAGTACCAGCGGGTCAATGAG,

(4) minimum average read quality, > 99.9% confidence (phred33 ≥ 30) per read,

(5) minimum single base pair quality, > 90% confidence (phred33 ≥ 10) per base pair,

(6) Nextera PE adaptor trimming, and

(7) 40 and 20 bp were excluded from left and right ends of the read, respectively, as these end parts showed sequencing errors in control samples (n = 8) as well as primer site. Thus, the mutation was quantified within a window of 81 bp on the left and 70 bp on the right from the canonical cleavage site, between third and fourth nucleotide upstream of preceding PAM sequence. The wide window was set to

maximise the detection of targeted mutation as some changes caused by CRISPR/Cas9 editing can appear away from the canonical cutting site as shown in Chapter 4 by Sanger sequencing.

Both graphical report and plain text output files were generated, and a text output, "Alleles_frequency_table.txt", was used for further integrative analyses using the Microsoft Excel program. Each mutation rate was calculated by the following formula:

- Mutation rate (%) = 100 – the percentage of NGS reads of WT

In addition, frame-shift mutation rate was calculated by the following formula:

- Frame-shift mutation rate (%) = frame-shift mutation reads (No.) / ∑total reads (No.) × 100

In addition, text outputs from CRISPResso, "effect_vector_deletion_NHEJ", "effect_vector_insertion_NHEJ" and "effect_vector_substitution_NHEJ" were used for distribution analysis of deletion, insertion, and substitution in relation to the reference amplicon, respectively, and "indel_histogram" was used for frequency analysis of indel size. Then, the frequency of each mutated allele among NGS reads was compared to the result from the fragment analysis. For clarity and visualisation purposes, alleles with a frequency of $\geq 0.5\%$ (*circa* 50 reads per individual) within each sample were designated as 'representative sequences' and used for mutation rate calculation, comparison with fragment analysis and indel size difference between mutant phenotypes. Alleles with a frequency of $> 0.01\%$ were designated as 'total sequences' and subjected to the analysis of the frequency of indel size and indel distribution.

6.2.3 Statistics

Statistical analysis was performed using Minitab 17 (Minitab Inc., State College, USA). Data are presented as mean ± SEM. Significant differences between group mean were tested by one-way ANOVA, followed by Tukey's HSD test ($p < 0.05$) for arbitrary gene modification rate by T7E1, number of different size fragments detected by fragment analysis, frame-shift mutation rate, number of different alleles and different size fragment analysed by NGS, proportion of indel size between phenotype groups analysed by fragment analysis and NGS, detected fragment number between fragment analysis and NGS, positional differences in deletion, insertion or substitution frequencies between the different phenotype groups around the canonical cut site). All percentage data were arcsine transformed and normality and homogeneity of variance were confirmed through

examination of the model residuals and fits and Levene's test, and where necessary data was further square root or $\log_{10}$ transformed.

In situations where the control samples were all registered as 0 or 1 (arbitrary gene modification rate by T7E1, and mutation rate, frame-shift mutation rate, number of different alleles and different size fragment analysed by NGS), it was excluded for statistical analysis. Significant differences between group mean which did not meet the normality of variance for arbitrary gene modification rates by RGEN and fragment analysis and mutation rate by NGS were assessed by Kruskal-Wallis test, followed by Mann-Whitney test ($p < 0.05$). The linear regression was performed by using GraphPad Prism v7.03 (GraphPad Software, La Jolla, USA) between the mutation rate determined by NGS and the arbitrary gene modification rates calculated by T7E1, RGEN and Fragment analyses.

6.3 Results

6.3.1 T7E1 assay

All 52 mutant samples showed the presence of three DNA bands of the expected size (undigested band of 970 bp, and two digested bands by T7E1, of 719 and 251 bp), while all the control samples only possessed an undigested band of 970 bp (Fig. 6.3A). The average arbitrary gene modification rate calculated by the T7E1 assay was significantly higher in phenotype A ($52.7 \pm 2.3\%$) in comparison to phenotype B ($39.2 \pm 2.7\%$) and C ($43.0 \pm 2.0\%$), and the control showed $0 \pm 0\%$ (Fig. 6.3B).

During the preparation of the DNA samples prior to the T7E1 assay, it was also noted that 19 out of 52 mutant samples (36.5%) showed additional bands of unexpected sizes on the gel electrophoresis of purified PCR products using *piwil2* sgRNA2_2 primer pair (Table 6.1&6.3 and Fig. 6.4). All individuals contained deletions of the size of -90 to -670bp while 2 individuals also contained putative insertions of 530bp and 580bp (Table 6.3, Fig. 6.4B). The putative large deletion fragment from a single sample (individual M1) was Sanger sequenced which confirmed a large deletion of 637 bp upstream adjacent to the PAM sequence in mutant 1 (Fig. 6.5).

(A)
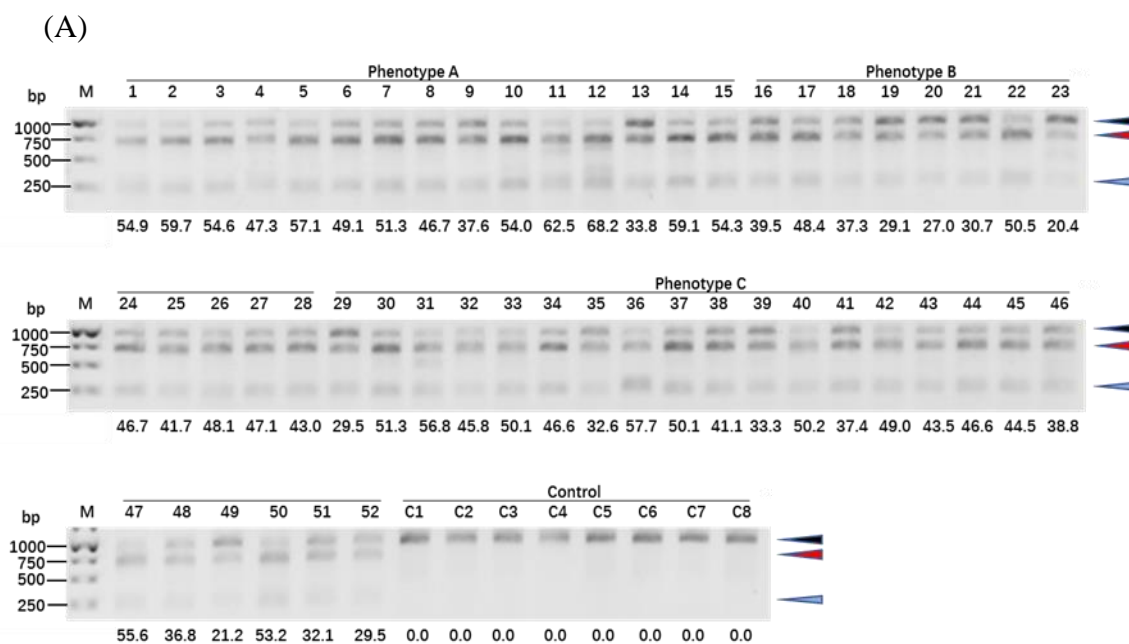
(B)
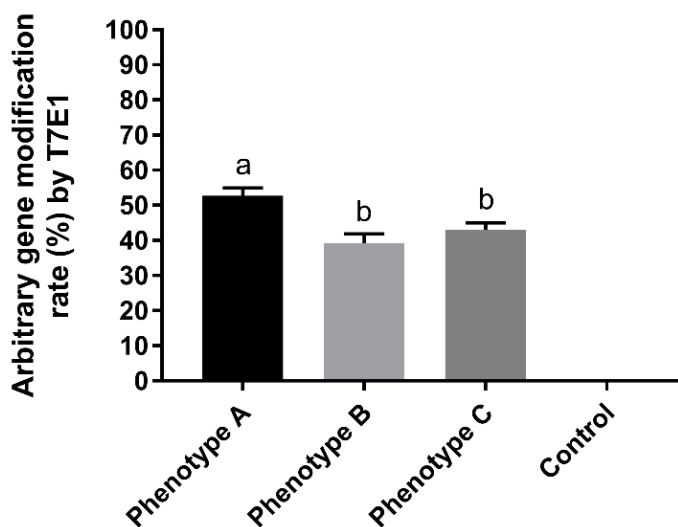
Figure 6.3. Genotyping of *piwil2* KO founders in phenotype A, B and C as well as WT control by T7E1 assay. (A) Arbitrary gene modification rate assessed by T7E1 assay is indicated below each sample. Phenotype A, sample number 1 - 15; phenotype B, sample number 16 - 28; phenotype C, sample number 29 - 52; and control, sample number C1-8. One uncleaved and two cleaved bands are indicated by three arrowheads on the right. (B) The average arbitrary gene modification rate of phenotype A ($n = 15$), B ($n = 13$), C ($n = 24$) and control ($n = 8$) by T7E1 assay. Data are shown as mean $\pm$ SEM. Superscripts denote a statistically significant difference between phenotypic groups ($p < 0.05$).

Table 6.3. Size and proportion of putative large indels (bp) shown in purified PCR products from *piwil2* sgRNA2_2 primer pair measured by GeneTool (Syngene, Cambridge, UK). The gel electrophoresis image of these samples is shown in Figure 6.4. 19 mutants out of 52 mutants (36.5%) showed putative large deletion (450.5 ± 37.4 bp, $n = 20$) or insertion (530 and 580 bp, $n = 2$) bands

| Sample No. | No. 1 | | No. 2 | | No. 3 | |
	Indel size	%	Indel size	%	Indel size	%
M01	-630	10.8				
M03	-350	1.8				
M04	-620	10.8				
M10	-540	1.3				
M11	530	6.1	-350	12.8		
M12	580	3.9	-300	7.7	-540	5.3
M13	-350	0.6				
M23	-90	8.4				
M24	-350	0.4				
M29	-670	4.5				
M31	-220	13.5				
M33	-470	8.3				
M34	-630	1.3				
M35	-630	0.5				
M36	-670	30.4				
M40	-390	1.2				
M41	-530	1.3				
M43	-420	5.6				
M46	-260	6.9				

(A)
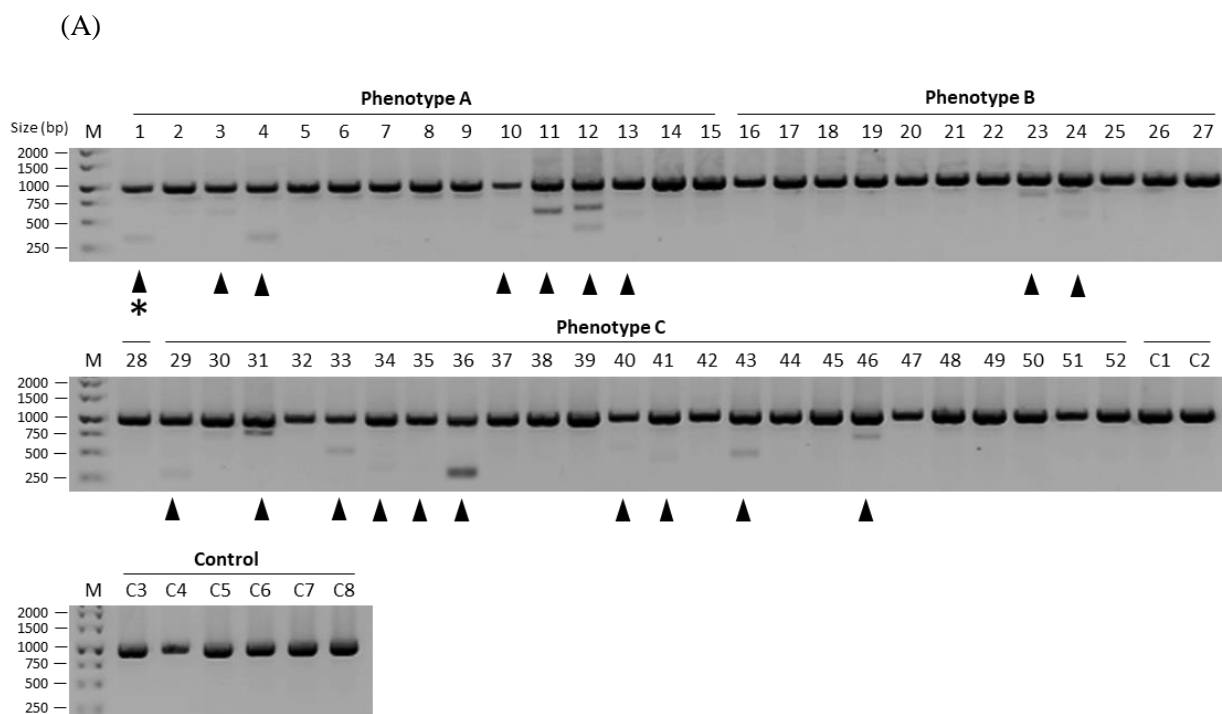
(B)

Figure 6.4. Putative large indels shown in purified PCR products from *piwil2* sgRNA2_2 primer pair (WT size is 970 bp). (A) Gel electrophoresis image of 52 mutants and controls. Arrowheads indicate the samples with large indels detected by GeneTool (Syngene, Cambridge, UK). The size and the proportional band intensity are shown in Table 6.3. Asterisk (*) indicates the sample which putative indel band was sequenced by Sanger sequencing. (B) Size and proportion of the putative large indels.

```
Control caaacaacaaaacacacacacacaaagcagaaggcaggcctctcatctttaacagcagcactgtgttt
Mutant1 caaacaacaaaacacacacacacaaagcagaaggcaggcctctcatct-------------------

Control acattaatgcattttgtaatcatttctcacatttataaaatgctgcctcagaaacaaactgacatcct
Mutant1 ------------------------------------------------------------------

Control acaatccttccaattggaatcagaccacttgttgttatccgcccgccctggactaaccaatctttctg
Mutant1 ------------------------------------------------------------------

Control agtgagggtgtgatccacgacagttccaggaggggggagtgccaaagctgtttgcagcccaggagtaca
Mutant1 ------------------------------------------------------------------

Control gagttgtgctaatgcgcttttgtaccacgatgaagaccagcttgggctcgtagctggggaaagtctca
Mutant1 ------------------------------------------------------------------

Control aagcacttgatcaactgtgggatctcgtactgctccaccatttttaagctggccgtctgagacaccatc
Mutant1 ------------------------------------------------------------------

Control tctgtacaccacaatcttttctggcaagttgtggtttacctgttagacatgggtaatcaaagtaaagt
Mutant1 ------------------------------------------------------------------

Control tatacaaatttaattcagtaaatgaaaagtgcattaaacccagaaacaaactaaactgttattttttat
Mutant1 ------------------------------------------------------------------

Control gacaatgcattcactgataggtgagaattaggtgtgggtttacagaatcacaacacatttctgaaatgc
Mutant1 ------------------------------------------------------------------

Control atttgtacctcatagtacttctgcagtgcagccagtaagcaaaccctgaagcc acggatcagttcctc
Mutant1 ------------------------------------------------------------------

Control attgg ggg tctggaaaattactctggagtaccagcgggtcaatgagctggaagggggacaacacactg
Mutant1 ----- ggg tctggaaaattactctggagtaccagcgggtcaatgagctggaagggggacaacacactg

Control ctggtaggactcatgcattgtgcatctcaccaattcaaggtcaaataatttctcatttgtgcactcat
Mutant1 ctggtaggactcatgcattgtgcatctcaccaattcaaggtcaaataatttctcatttgtgcactcat

Control atacaaagacatgaaaagctttatttgtcaaggtgatgcaaaaatgcttgaaaaaaagtgagtcacag
Mutant1 atacaaagacatgaaaagctttatttgtcaaggtgatgcaaaaatgcttgaaaaaaagtgagtcacag

Control cattgctttgcattgtctagtaat
Mutant1 cattgctttgcattgtctagtaat
```

Figure 6.5. Sequence of the putative large deletion PCR fragment in *piwil2* mutant No. 1 aligned with the control. The putative large deletion PCR fragment of mutant 1 was gel extracted, amplified by 2nd PCR and sequenced by Sanger sequencing in both directions. The contigs of mutant 1 and the control were aligned by MAFFT (https://www.genome.jp/tools-bin/mafft). The target area of *piwil2* sgRNA2 is boxed and PAM sequence is in bold. Dash means deletion and it possesses a large deletion of 637 bp upstream the PAM sequence.

6.3.2 RGEN assay

None of the 52 mutants showed complete digestion by RGEN, while control samples showed almost or completely digested bands by RGEN (Fig. 6.6A). The average arbitrary gene modification rate assessed by RGENs was significantly higher in phenotype A (91.7 ± 1.9%) than C (81.1 ± 6.7%) ($p < 0.05$), while phenotype B (82.6 ± 2.9%) was not significantly different from either phenotype A or B ($p > 0.05$) (Fig. 6.6B). All of the three mutant groups showed significantly higher arbitrary mutation rate than control (7.7 ± 2.3%) ($p < 0.05$) (Fig. 6.6B).

(A)
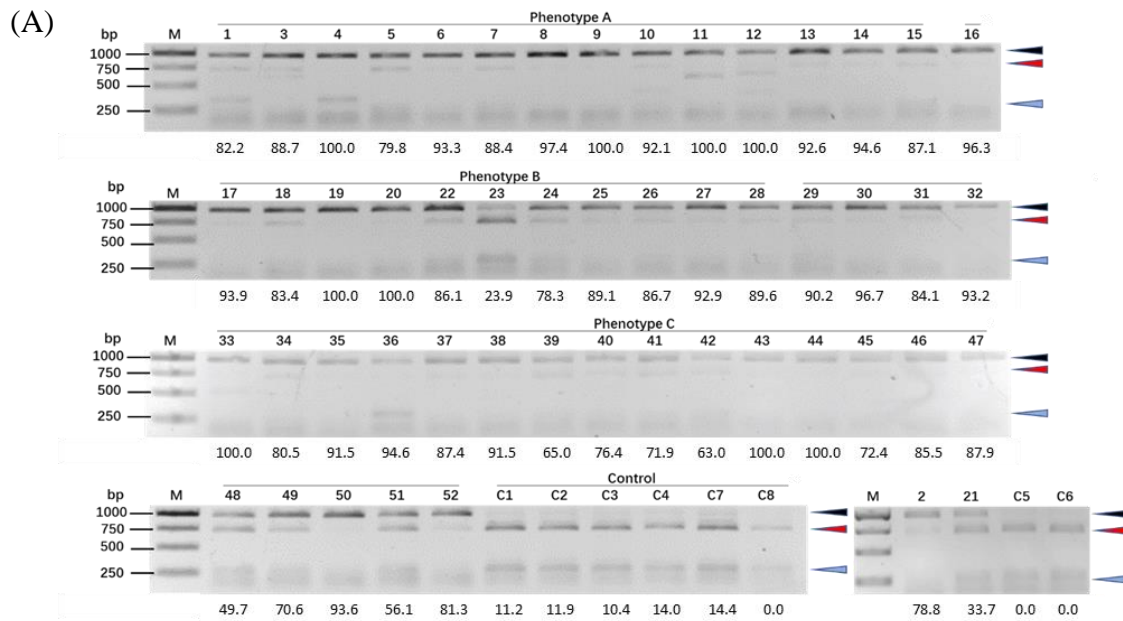
(B)
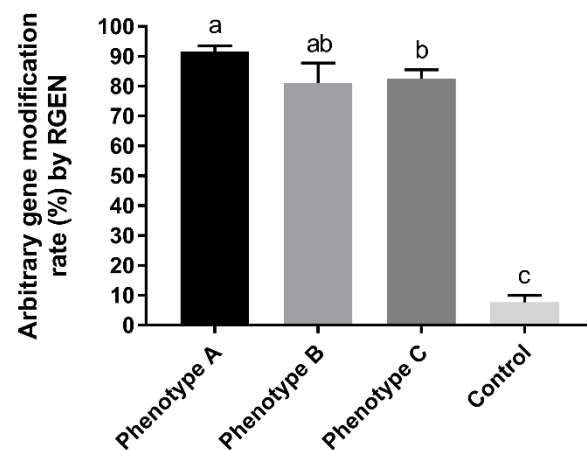
Figure 6.6. Genotyping of *piwil2* KO founders in phenotype A, B and C as well as WT control by RGEN assay. (A) Arbitrary gene modification rate assessed by RGEN assay is indicated below each sample. Phenotype A, sample No. 1-15; phenotype B, sample No. 16 - 28; phenotype C, sample No. 29 - 52; and control, sample No. C1 - 8. One uncleaved and two cleaved bands are indicated by three arrowheads on the right. (B) The average arbitrary gene modification rate of phenotype A ($n = 15$), B ($n = 13$), C ($n = 24$) and control ($n = 8$) by RGEN assay. Data are shown as treatment mean $\pm$ SEM. Superscripts denote statistically significant difference between groups ($p < 0.05$).

6.3.3 HRMA

All the WT control melt curves were tightly grouped (average of $85.23 \pm 0.01°C$), and all the mutant samples were clearly distinguishable from the WT control melt curves (Fig. 6.7A). Although the boundaries of the melt curves were not distinct between mutant phenotypes, the majority of phenotype A mutants tend to be found in the upper range of the melt curve cluster, *i.e.* earlier melting point (Fig. 6.7B), while phenotype C was in the middle (Fig. 6.7D) and phenotype B in the lower range (Fig. 6.7C). All WT control samples had only one Tm peak ranging from 85.20 to 85.24°C, while mutants displayed between 2 or 3 identifiable Tm peaks ranging from 80.8 to 85.7°C (Fig. 6.8). The distribution of Tm peaks of each phenotype showed three clusters of Tm peaks (Fig. 6.8), where first, second and third peak clusters were distributed between 80.8 – 81.8°C, 82.0 – 84.2°C and 84.4 – 85.7°C, respectively, and there was no obvious difference in the distribution of Tm peak between phenotype groups.

The shape of the HRMA melt curves were classified into four types: type I with two to three major peaks with one of them located in the first Tm cluster (80.8 – 81.8°C) (Fig. 6.9A); type II with a dominant peak in the second Tm cluster (82.0 – 84.2°C) (Fig. 6.9B); type III with a dominant peak in the third Tm cluster (84.4 – 85.7°C) (Fig. 6.9C); and control with a single dominant peak at 85.2°C (Fig. 6.9D). In phenotype A group, there were 60.0, 20.0 and 20.0% of melt curve type I, II and III, respectively, while it was 15.4, 15.4 and 69.2% in phenotype B and 16.7, 25.0 and 58.3% in phenotype C (Fig. 6.9E).

(A)

(B)

(C)

(D)
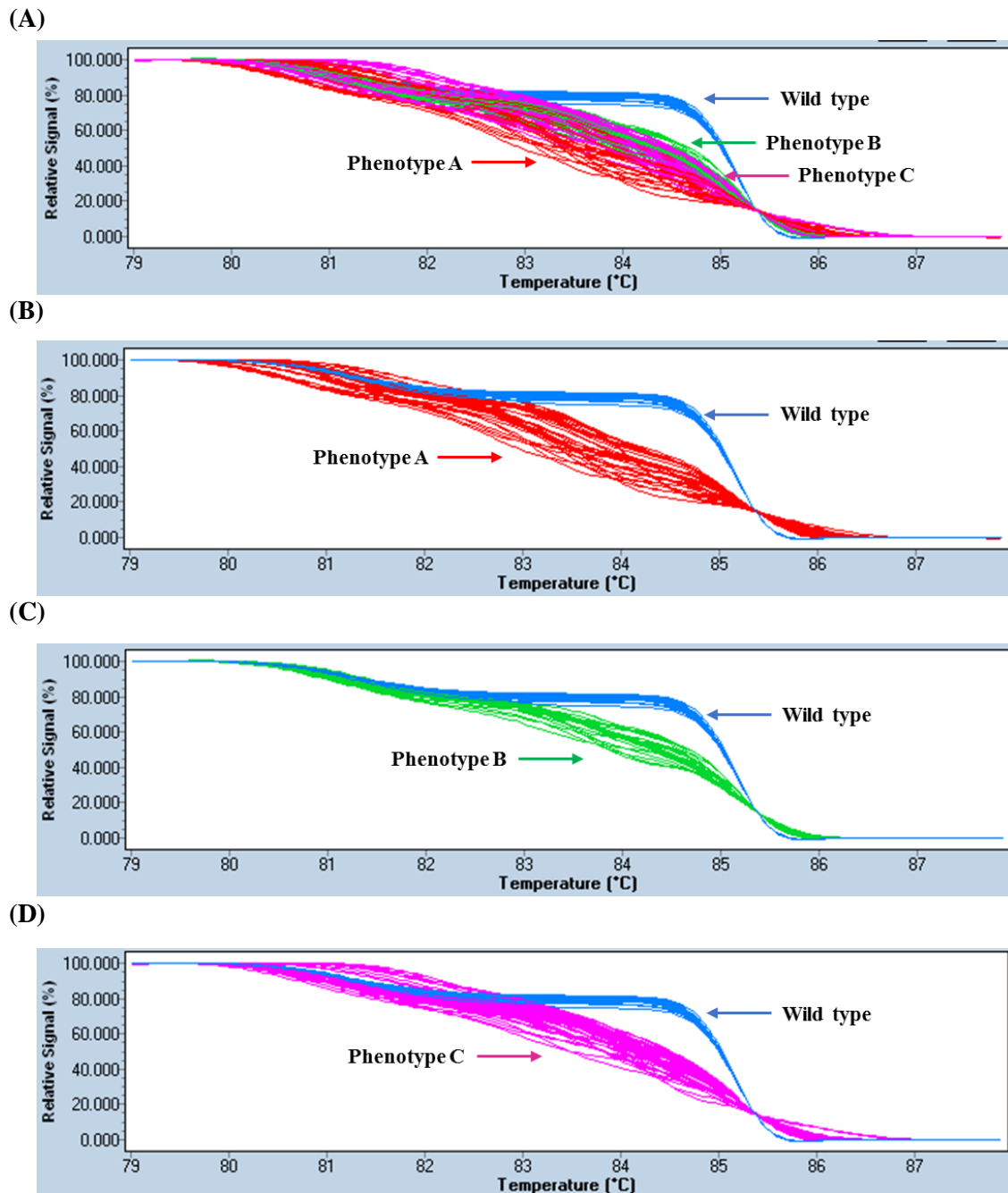
Figure 6.7. HRMA of *piwil2* mutants and control. (A) Melt curves of WT control and mutant phenotype A, B and C. (B) Melt curves of phenotype A, red ($n = 15$), (C) phenotype B, green ($n = 13$) and (D) phenotype C, purple ($n = 24$), clearly distinguishable from the WT control, blue ($n = 8$).
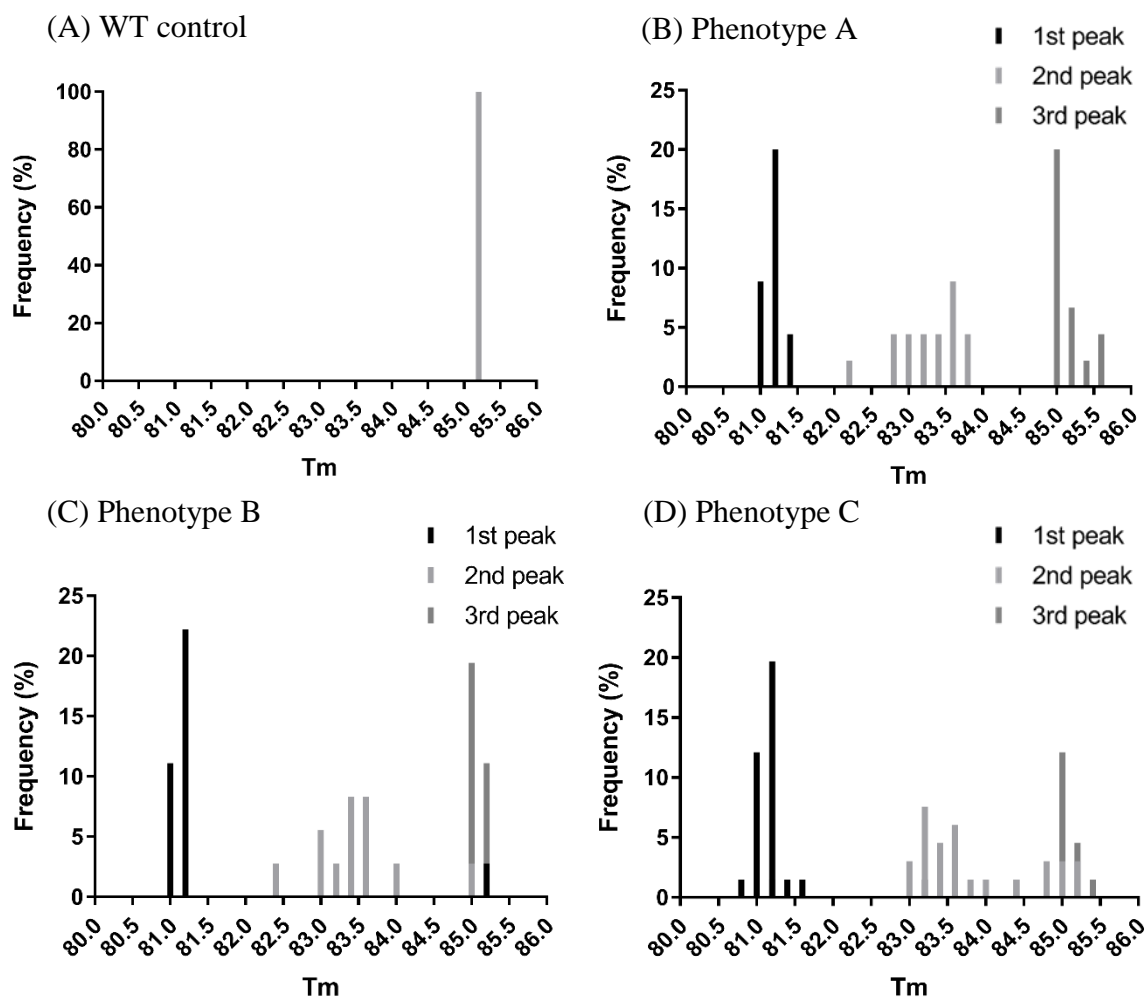
Figure 6.8. Distribution and frequency of Tm peaks as identified by HRMA. (A) WT control ($n = 8$), (B) phenotype A ($n = 15$), (C) phenotype B ($n = 13$) and (D) phenotype C ($n = 24$).

(A) Melt curve type I

(B) Melt curve type II

(C) Melt curve type III

(D) Control melt curve

(E)
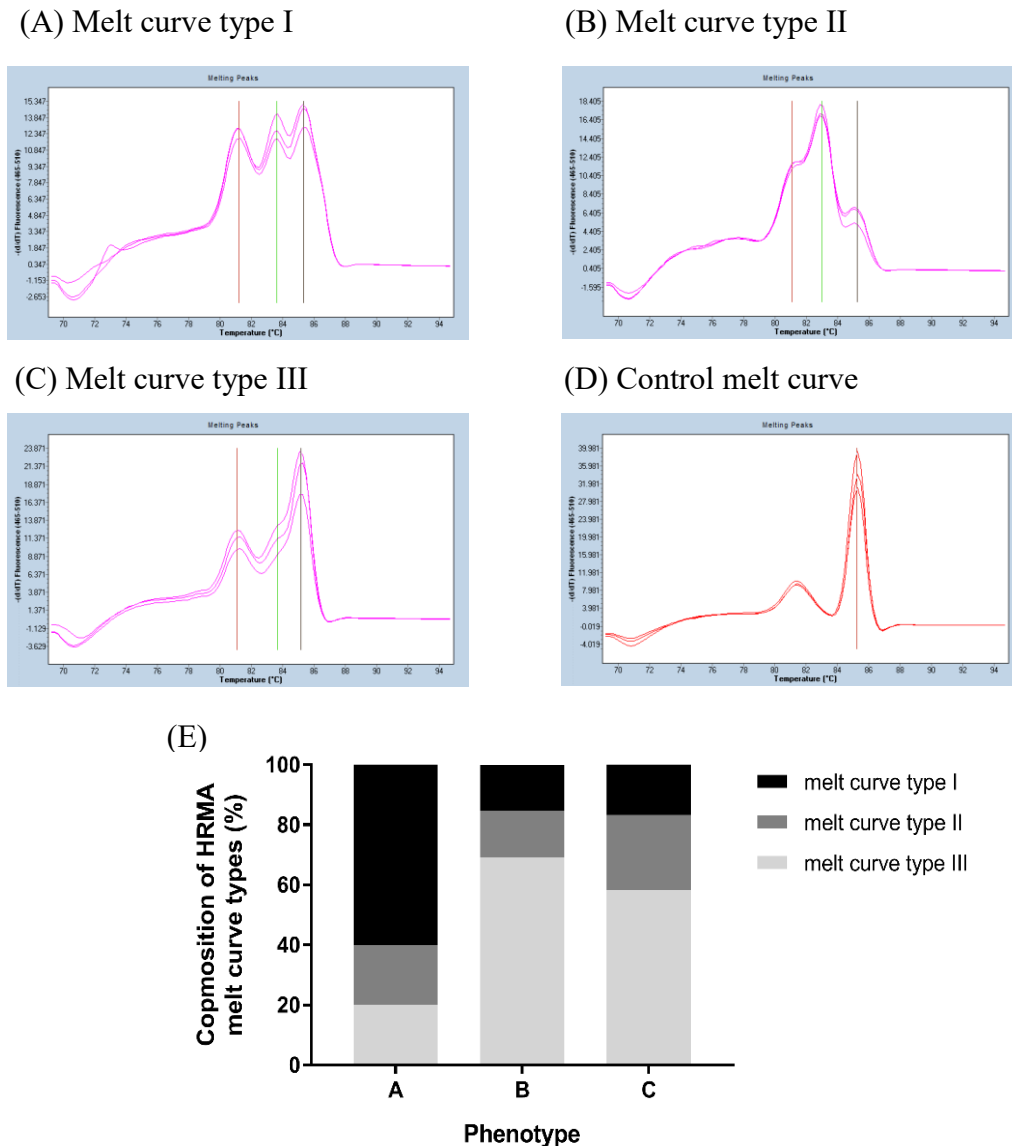
Figure 6.9. Representative HRMA melt curve type and the composition of the melt curve types in each phenotype of *piwil2* mutants. (A-D) Representative shapes of HRMA melt curve types. (A) HRMA melt curve type I, two to three dominant peaks with one of them ranging between 80.8 – 81.8°C; (B) type II, a dominant peak between 82.0 – 84.2°C; (C) type III, a dominant peak between 84.4 – 85.7°C; and (D) control, one dominant peak at 85.2°C. (E) Composition of HRMA melt curve types in each phenotype of *piwil2* mutants. Data are presented as percentage (%) of the melt curve type in each phenotype.

6.3.4 Fragment analysis

All the WT control showed a dominant peak at 228 nt (Fig. 6.10A); however, six out of eight controls had additional minor peaks of low abundance ($5.81 \pm 0.23\%$ of total measured fragments). 17 out of 52 mutant samples did not have measurable WT size fragments (Fig. 6.10B), while the rest of mutants had a measurable WT size fragment (Fig. 6.10C).

The average arbitrary gene modification rate analysed by fragment analysis were comparable ($p > 0.05$) between phenotype A ($91.7 \pm 1.9\%$), B ($81.1 \pm 6.7\%$) and C ($82.6 \pm 2.9\%$), while all mutant phenotypes were significantly higher ($p < 0.05$) than WT control ($7.7 \pm 2.3\%$) (Fig. 6.11A). In addition, there was no significant difference in the average number of recorded fragments in relation to the mutant phenotypes (Phenotype A, B and C, 17 ± 1 peaks, 13 ± 2 peaks and 15 ± 1 peaks, respectively), while all mutant phenotypes were significantly higher ($p < 0.05$) than WT (1.0 ± 0.3 peaks) (Fig. 6.11B).

The mean proportion of deletions greater than 5 bp was significantly higher in phenotype A than in phenotype B and C, while there was no difference in the proportion of any other identified indel fragments including WT and mutant sequences without changes in absolute length (WT length mutated sequences) in relation to the three phenotype groups (Fig. 6.12A).

(A) WT (control 7)

(B) Mutant without WT length fragment (mutant 11, phenotype A)

(C) Mutant with WT length fragment (mutant 21, phenotype B)
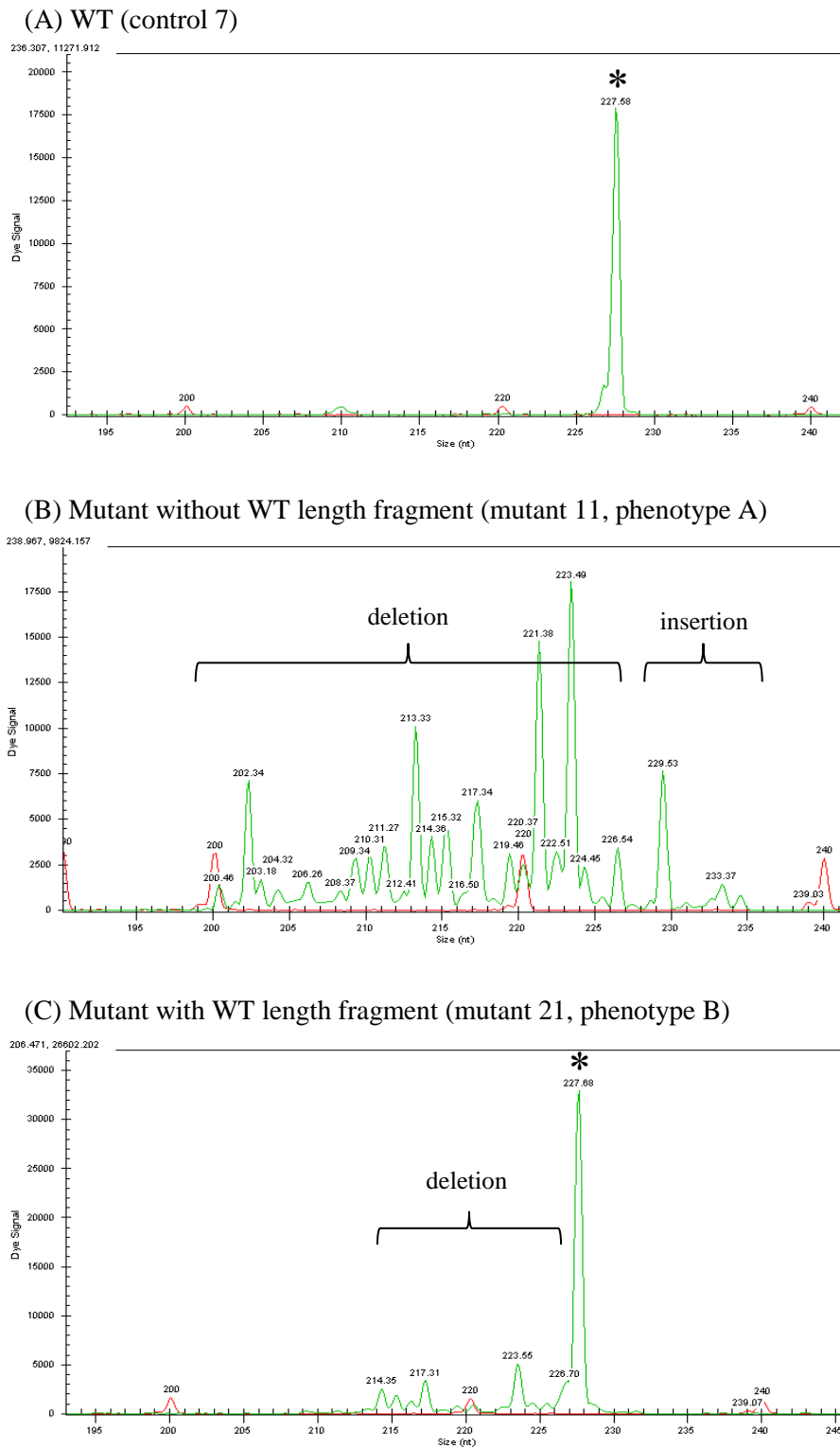
Figure 6.10. Representative results of fragment analysis. (A) WT control, (B) *piwil2* mutant which does not possess WT length fragment and (C) *piwil2* mutant which possesses WT length fragment. WT length fragment is denoted by asterisk (*). The peaks of size marker are present at 200, 220 and 240 bp with red line.

(A)

(B)
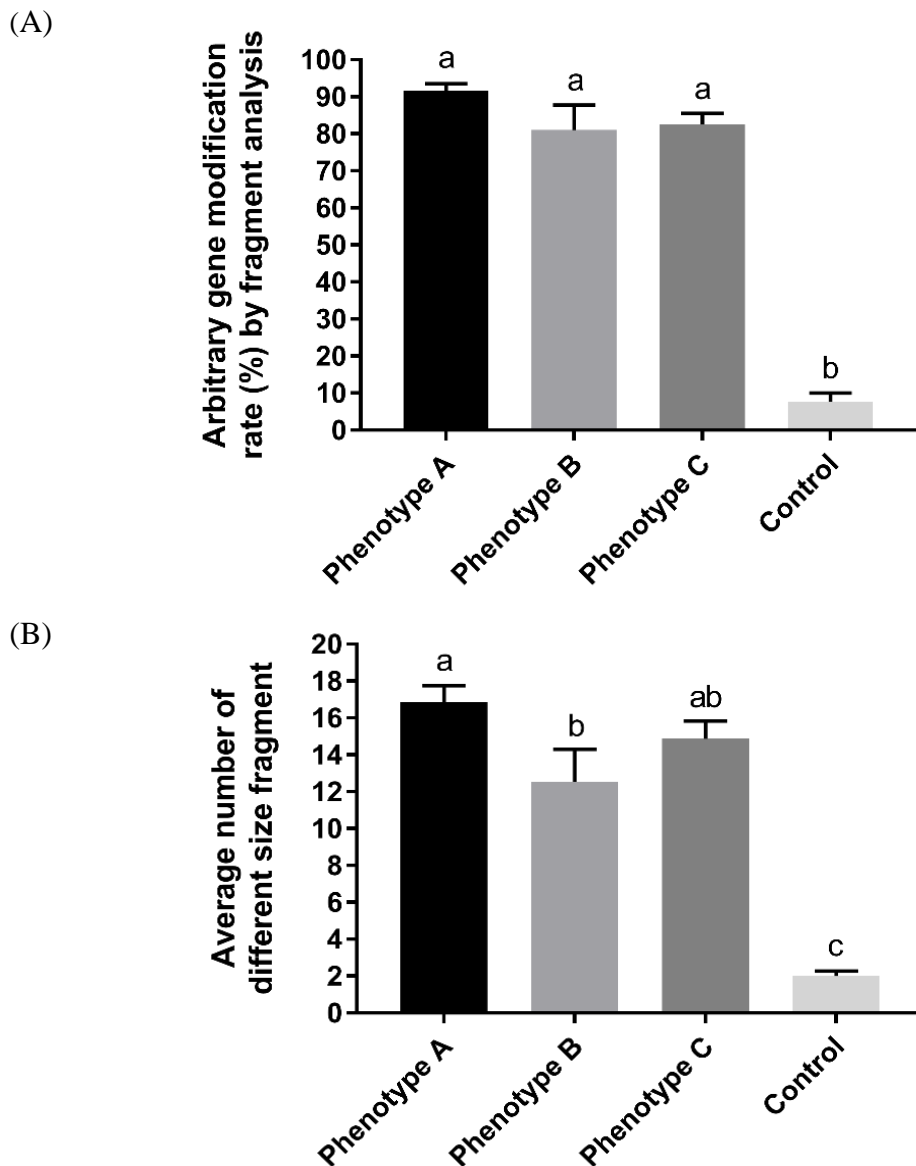
Figure 6.11. Genotyping of *piwil2* KO larvae grouped by observed phenotype A, B and C as well as WT control by fragment analysis. (A) The average arbitrary gene modification rate of phenotype A ($n = 15$), B ($n = 13$), C ($n = 24$) and control ($n = 8$) assessed by fragment analysis. (B) The average number of fragments detected by fragment analysis in phenotype A, B, C and control. Data are shown as treatment mean $\pm$ SEM. Superscripts denote statistically significant difference between groups ($p < 0.05$).

(A) Fragment analysis
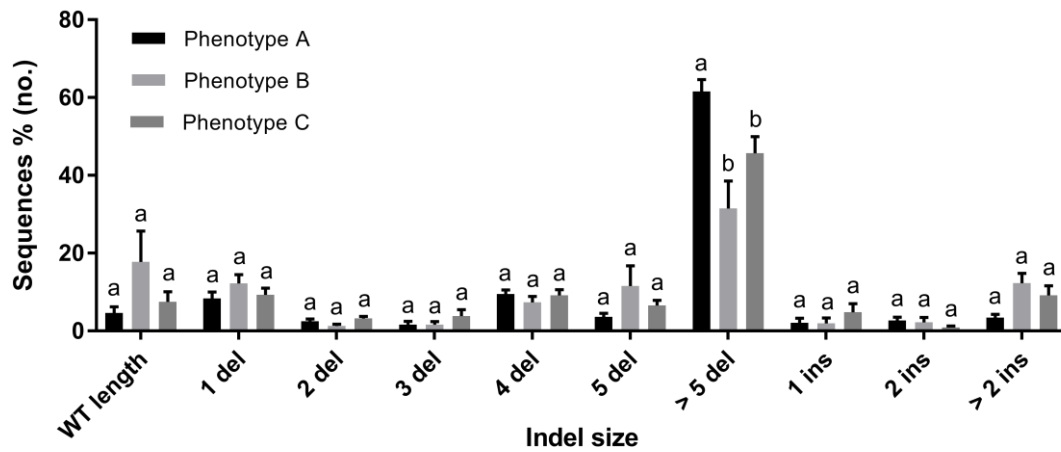
(B) NGS
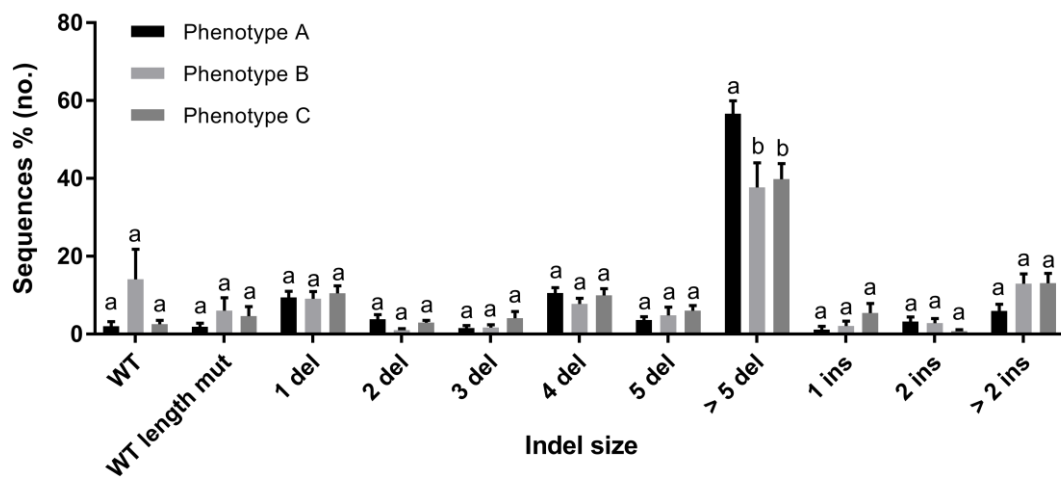
Figure 6.12. Average proportion of indel size in phenotype A, B and C mutants. (A) The different distribution of indel sizes in phenotype A ($n = 15$), B ($n = 13$) and C ($n = 24$) was assessed by fragment analysis (A) or NGS using the representative sequences (B). Data are shown as mean ± SEM. Superscripts denote statistically significant difference between phenotypic groups at each indel size ($p < 0.05$). WT, wild type sequence; WT length, WT and WT length mutant; WT length mut, WT length mutant; del, deletion; ins, insertion.

6.3.5 NGS

The average mutation rates determined by NGS were not significantly different between mutant phenotype groups, showing 97.9 ± 1.2, 85.9 ± 7.7 and 97.4 ± 1.0% in phenotype A, B and C, respectively, while the WT control was 0% (Fig. 6.13A). The frame-shift mutation rates were 78.3 ± 1.8, 73.8 ± 4.0 and 74.0 ± 4.1% in phenotype A, B and C, respectively, and they were comparable among mutant phenotype groups (Fig. 6.13B). The potential splice site mutation rates were 0.5 ± 0.3, 0.3 ± 0.2 and 0.4 ± 0.2% in phenotype A, B and C, respectively, which were not significantly different between mutant phenotype groups. The average numbers of different alleles detected were 28 ± 2, 19 ± 2 and 23 ± 1 per individual in phenotype A, B and C, respectively, and it was significantly higher in phenotype A than B and C (Fig. 6.13C). The average numbers of different size fragments detected by NGS were 20 ± 1, 15 ± 1 and 17 ± 1 in phenotype A, B and C, respectively, and it was significantly higher in phenotype A than B, but it was comparable between A and C or B and C (Fig. 6.13D).

6.3.6 Analysis of indel diversity as identified by NGS

Proportion of indel size between phenotype groups analysed by NGS showed similar result with fragment analysis, with a mean proportion of deletions greater than 5 bp significantly higher in phenotype A than in phenotype B and C. However, there was no difference in the proportion of any other identified indel fragments including WT and WT length mutated sequences (mutant sequences without changes in absolute length) in relation to the three phenotype groups (Fig. 6.12B).
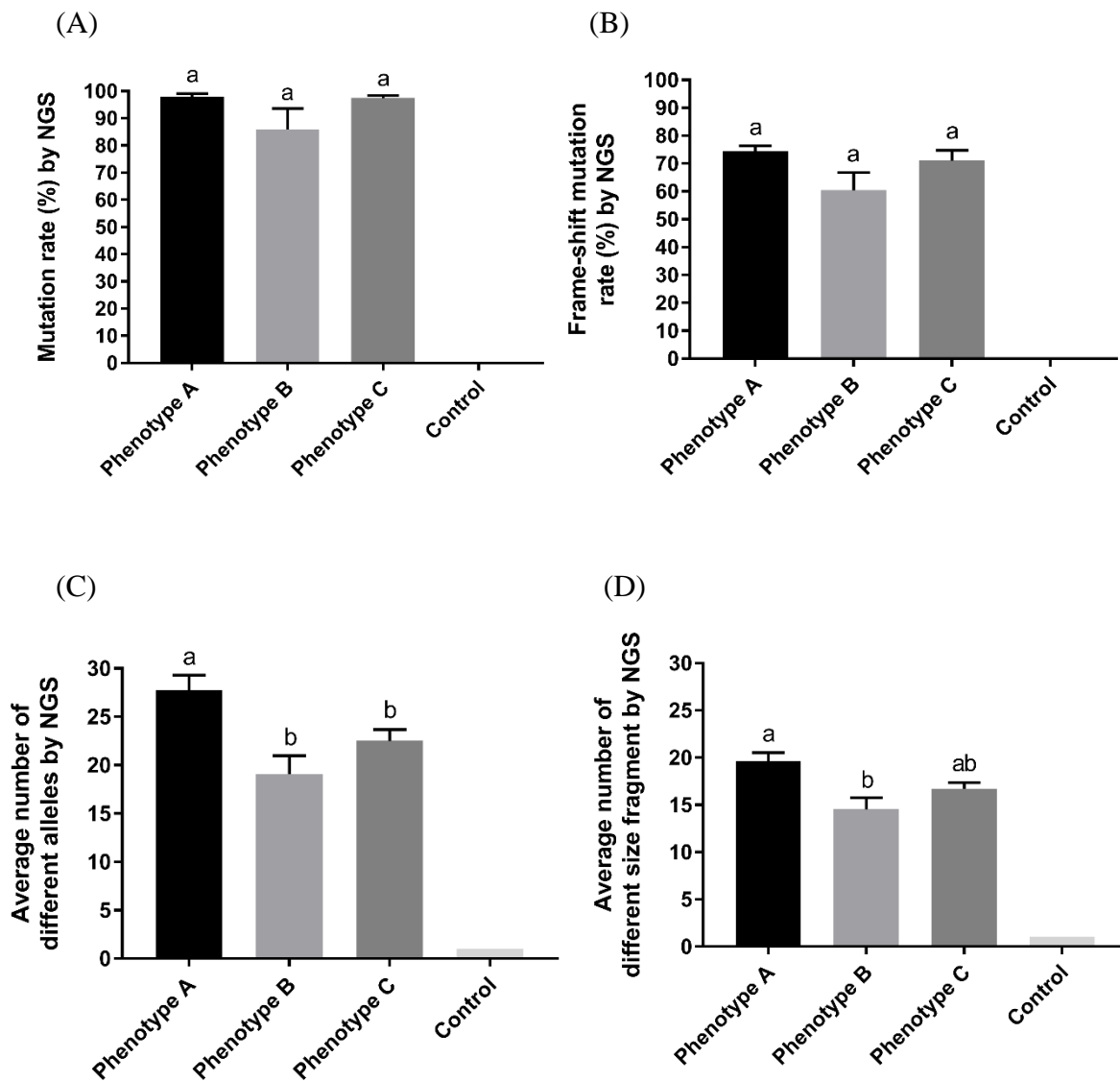
**Figure 6.13**. Genotyping of *piwil2* KO larvae grouped by observed phenotype A, B and C as well as WT control by NGS. (A) The average mutation rate, (B) the average frame-shift mutation rate, (C) the average number of different alleles per individual in the representative sequences and (D) the average number of different size fragments per individual in the representative sequences of phenotype A ($n = 15$), B ($n = 13$), C ($n = 24$) and control ($n = 8$) assessed by NGS. Data are shown as treatment mean ± SEM. Superscripts denote statistically significant difference between groups ($p < 0.05$).

The frequency of deletion events in the 52 mutants was highest at the predicted cleavage position (121 bp in reference to the WT amplicon position) with $75.5 \pm 3.4\%$ of sequences recording a deletion event at this location. The overall range of deletion activity (defined as locations where 1% or greater frequency of deletion event recorded) extended from 44 bp upstream to 17 bp downstream from the canonical cleavage site (77 – 138 bp in reference amplicon position) (Fig. 6.14A). There was a trend for deletion activity to be higher in phenotype A; however, it was only significantly higher than phenotype B at positions 114, 117 and 120-122 (Fig. 6.14B). The highest frequency of insertion was also shown in the cleavage site with $12.6 \pm 1.7\%$ of sequences recording an insertion event at this location (Fig. 6.15A). The distribution of insertion activity was narrower than deletion, showing the window of 7 bp upstream and 4 bp downstream from the cleavage site (114 – 125 bp in reference amplicon position) where insertion activity was observed in at least 1% of sequence reads (Fig. 6.15A). There was a trend for insertion activity to appear lowest in Phenotype A; however, it was only significantly lower than phenotype C at position 122 (Fig. 6.15B). The highest frequency of substitution ($12.5 \pm 1.7\%$) was shown in the cleavage site (Fig. 6.16B) like the insertion (Fig. 6.15B). The distribution of substitution activity ranged from 3 bp upstream to 7 bp downstream from the cleavage site (118 – 128 bp in reference amplicon position) (Fig. 16A). There was no apparent difference in substitution activity between the phenotype groups (Fig. 6.16B).

Overall, the total frequency of indel fragments observed in relation to phenotype groupings was comparable between the three groupings (Fig. 6.17A) with the spread in indel size expressed as 5 – 95th percentile or 25 – 75th percentile being comparable between the three phenotype groupings (Table 6.4 & Fig. 6.17B).

(A) All 52 mutants
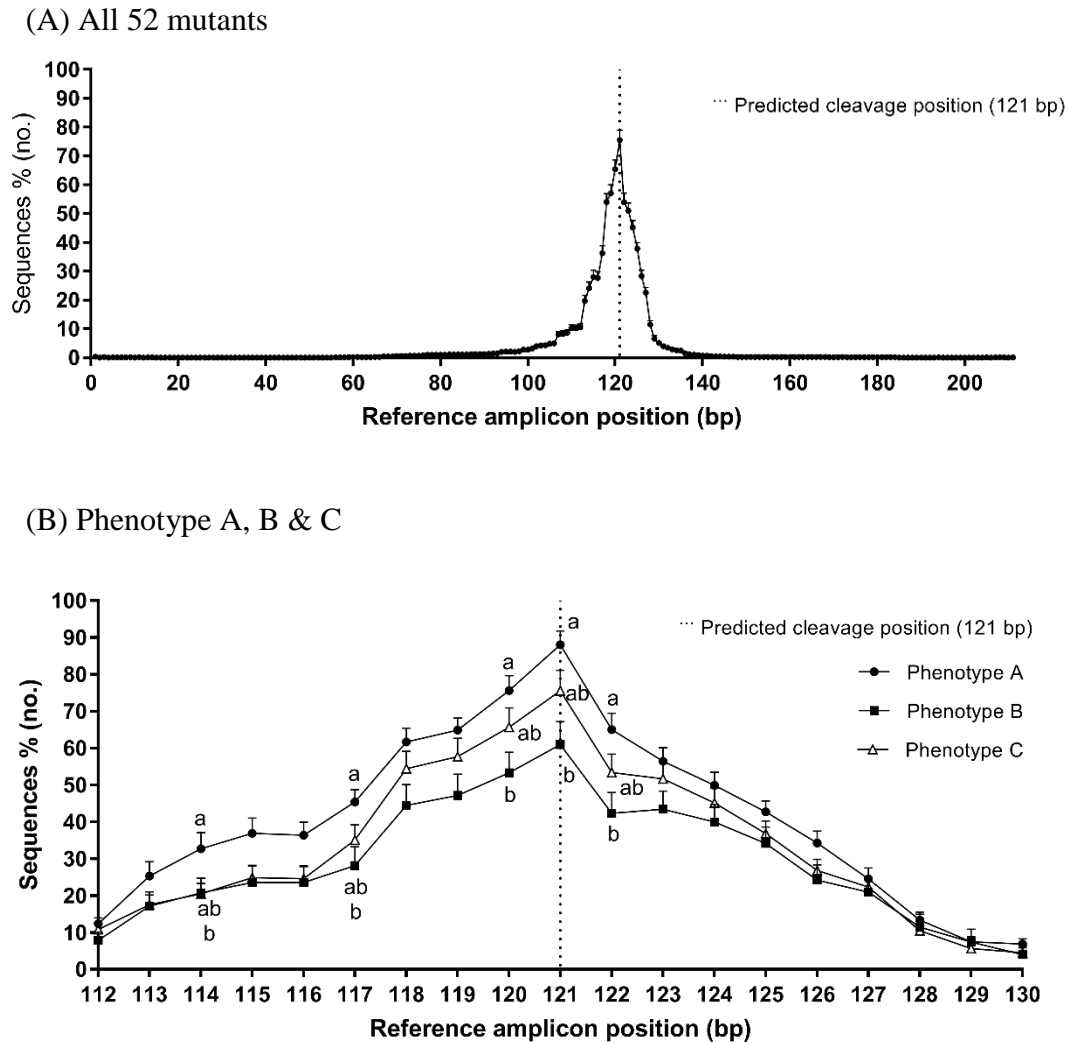
(B) Phenotype A, B & C

Figure 6.14. Deletion mutation position distribution. (A) the average of all 52 mutants and (B) the average of phenotype groups A ($n = 15$), B ($n = 13$) and C ($n = 24$). The predicted cleavage position is indicated by the dotted line at 121 bp in the total 211 bp amplicon. All sequences from the NGS results were used. Data are shown as mean ± SEM. Superscripts denote statistically significant difference between phenotype groups within the same position (bp) ($p < 0.05$).
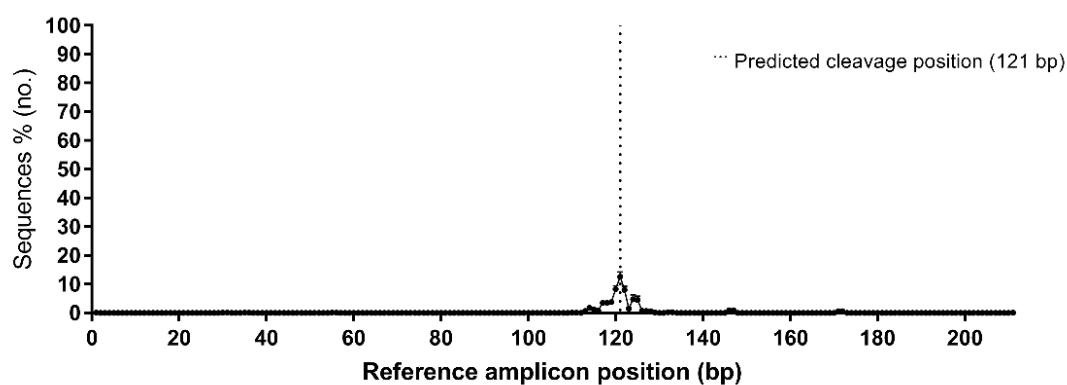
(A) All 52 mutants
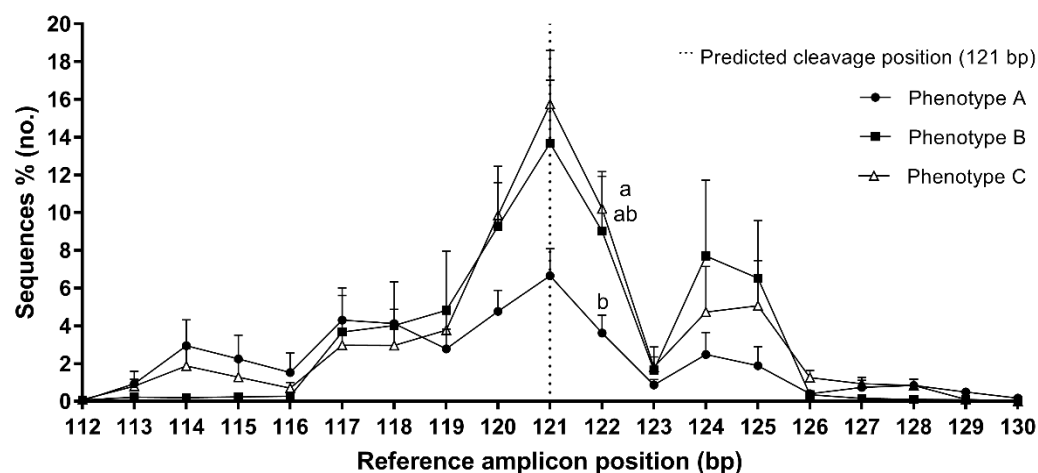
(B) Phenotype A, B & C

Figure 6.15. Insertion mutation position distribution. (A) the average of all 52 mutants and (B) the average of phenotype group A ($n = 15$), B ($n = 13$) and C ($n = 24$). The predicted cleavage position is indicated by the dotted line at 121 bp in the total 211 bp amplicon. All sequences from the NGS results were used. Data are shown as treatment mean ± SEM. Superscripts denote statistically significant difference between phenotype groups within the same position (bp) ($p < 0.05$).
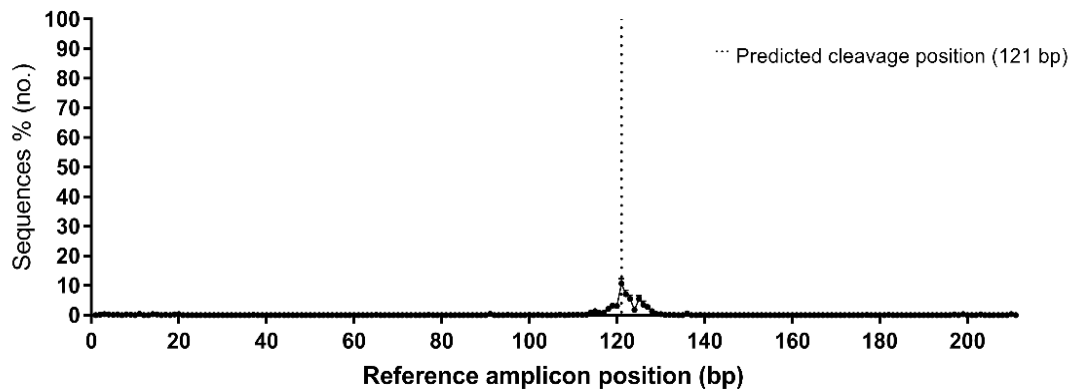
(A) All 52 mutants
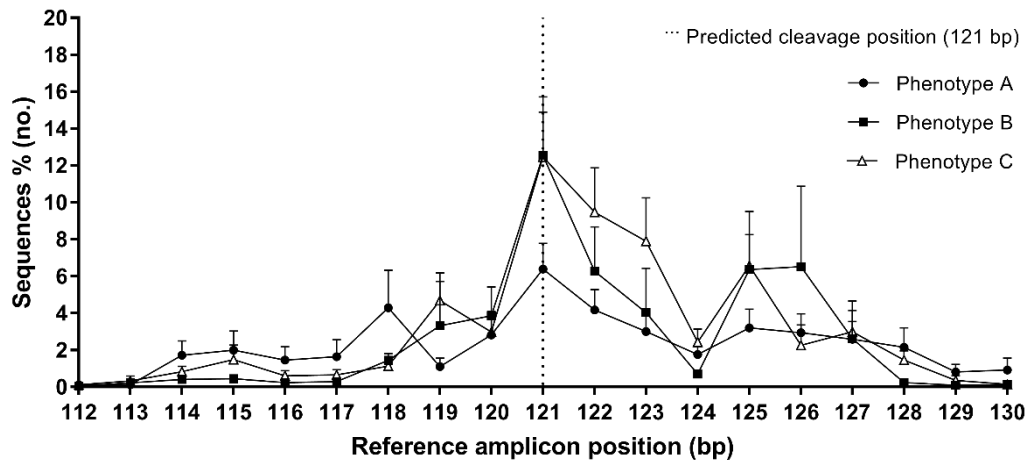
(B) Phenotype A, B & C

Figure 6.16. Substitution mutation position distribution. (A) the average of all 52 mutants and (B) the average of phenotype group A ($n = 15$), B ($n = 13$) and C ($n = 24$). The predicted cleavage position is indicated by the dotted line at 121 bp in the total 211 bp amplicon. All sequences from the NGS results were used. Data are shown as mean ± SEM.

(A) Frequency of indel size
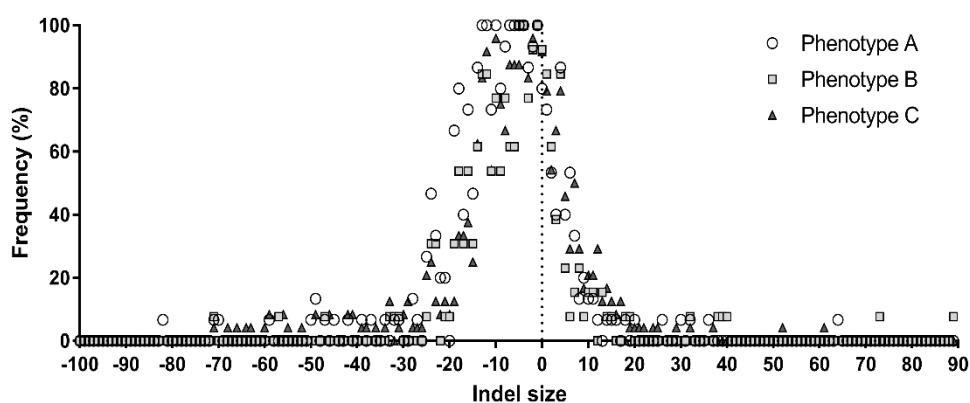
(B) Indel size (5th and 95th percentiles)
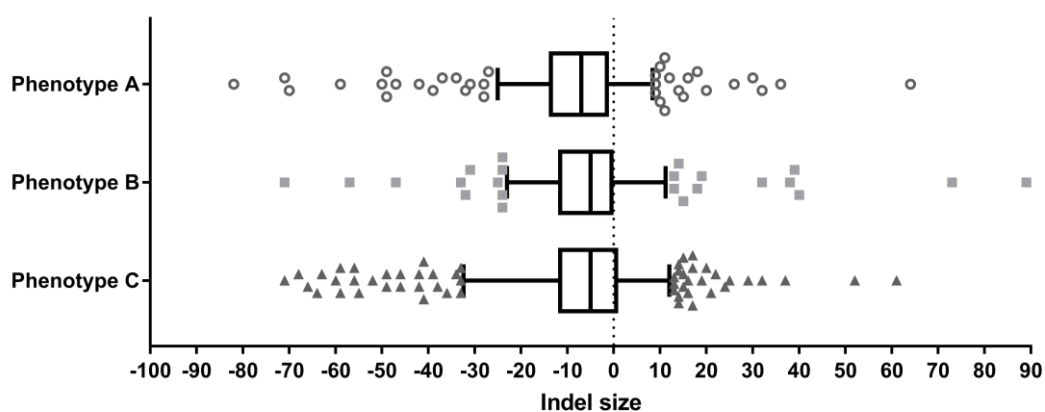
Figure 6.17. Frequency of indel size in each phenotype. (A) Each dot indicates the frequency of each indel size in each mutant phenotype groups. For example, 100% of -11 bp in phenotype A means all mutants in group A possess -11 bp of indel size regardless of reads % of the indel size of -11 bp in each individual. (B) the box shows 25th, 50th and 75th percentiles and whiskers show 5th and 95th percentiles. Indel sizes located outside of $5 – 95^{th}$ percentile are shown as dots.

Table 6.4. Percentiles (5th, 25th, 50th, 75th and 95th) of indel size distribution in each phenotype group

	5th	25th	50th	75th	95th
Phenotype A	-25	-14	-7	-1	8.4
Phenotype B	-23	-12	-5	0	11.2
Phenotype C	-32.4	-12	-5	1	12

6.3.7 Comparison of mutant screening methods

Regression analysis between mutation rates determined by NGS and the arbitrary gene modification rate assessed by T7E1, RGEN and fragment analysis revealed the highest correlation was with fragment analysis (R^2=0.72), a moderate correlation with RGEN (R^2=0.59) and a weak correlation with T7E1 (R^2=0.10) (Fig. 6.18). In general, genotype diversity identified by NGS and fragment analysis was similar (see typical examples in Fig. 6.19A-C); however, on closer consideration of the data, it was apparent that fragment analysis has a tendency to underestimate the number of fragments (as a proxy for alleles) both in comparison to fragment lengths and true alleles as recorded by NGS (Fig. 6.19D). This was primarily due to the tendency of fragment analysis to miss low abundance fragments ($< 2\%$) such as insertion fragments in Fig. 6.19B & C. In addition, plotting of the mutation rate calculated by fragment analysis and NGS revealed four apparent outliers where the gene modification rate by fragment analysis was underestimated in comparison to NGS (Fig. 6.19C). Further examination of the NGS output for these four cases revealed that these samples possessed a high proportion of mutant sequences (ranging from 12 to 51.8% of total sequences) which were modified without changes in absolute length (*i.e.* WT length mutated sequences) (Table 6.5). Such cases could not be differentiated from WT sequences by fragment analysis driving the apparent disparity in mutation rate calculation in these specific cases (Table 6.5).

(A)

(B)

(C)
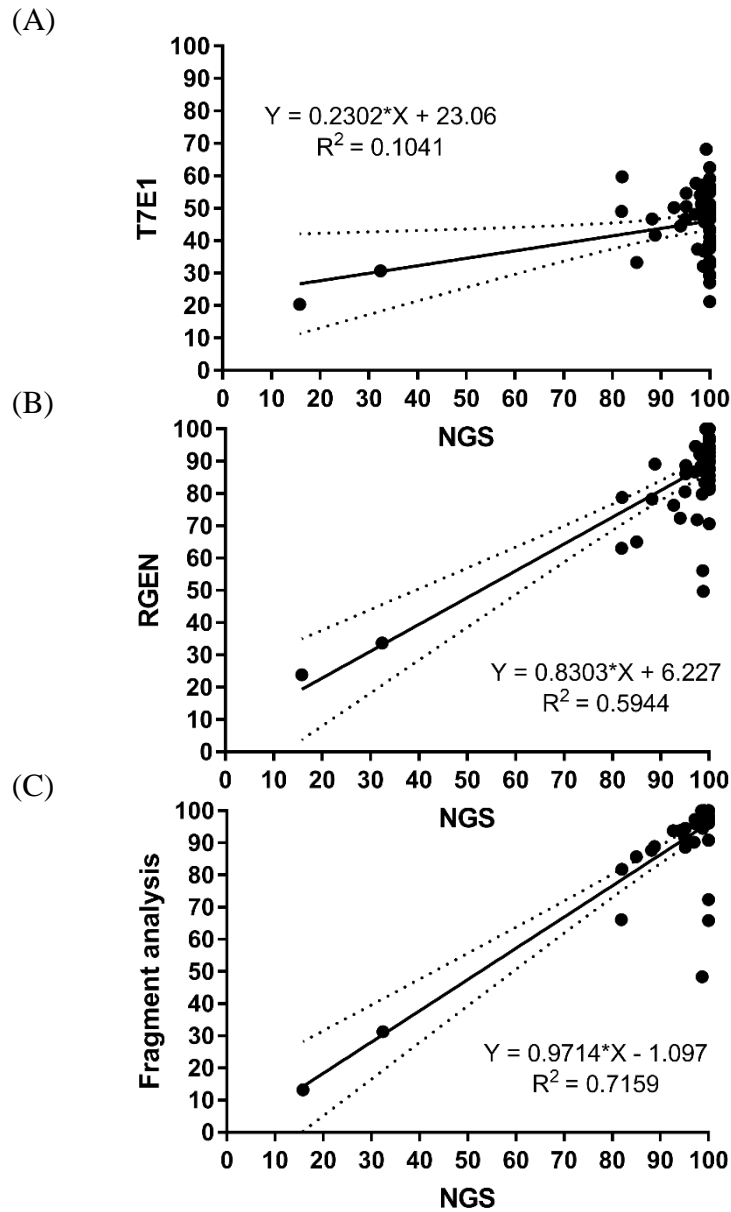
Figure 6.18. Scatter plot and linear regression between the mutation rate determined by NGS and the arbitrary mutation frequency calculated by (A) T7E1, (B) RGEN or (C) fragment analysis in 52 mutant samples. Each resulting regression is drawn with its confidence interval at 95% (dotted line).

Table 6.5. List of mutants which were outliers in regression analysis between fragment analysis and NGS shown in Fig. 6.18C. WT length mutated sequences are mutant sequences without changes in absolute length. del, deletion (bp); ins, insertion (bp); mut, mutation (bp); sub, substitution

	WT length fragment (%)	WT (%) in NGS	WT length mutated seq (%) in NGS (No. of seq)	Description of WT length mutated seq	WT + WT length mutants (%) in NGS
mutant 28	34.1	0.0	36.6 (2)	33.3% of 2 del, 2 ins and 2 mut 3.3% of 1 sub	36.6
mutant 42	33.9	18.1	12.0 (3)	7.9% of 2 mut 3.1% of 2 mut 1% of 3 del, 3 ins and 3 mut	30.1
mutant 51	51.6	1.3	51.8 (2)	42.4% of 2 mut 9.4% of 3 del, 3 ins and 2 mut	53.0
mutant 52	27.6	0.0	31.0	29.1% of 3 mut 1.9% of 1 del, 1 ins	31.0

(A) WT (control 7)
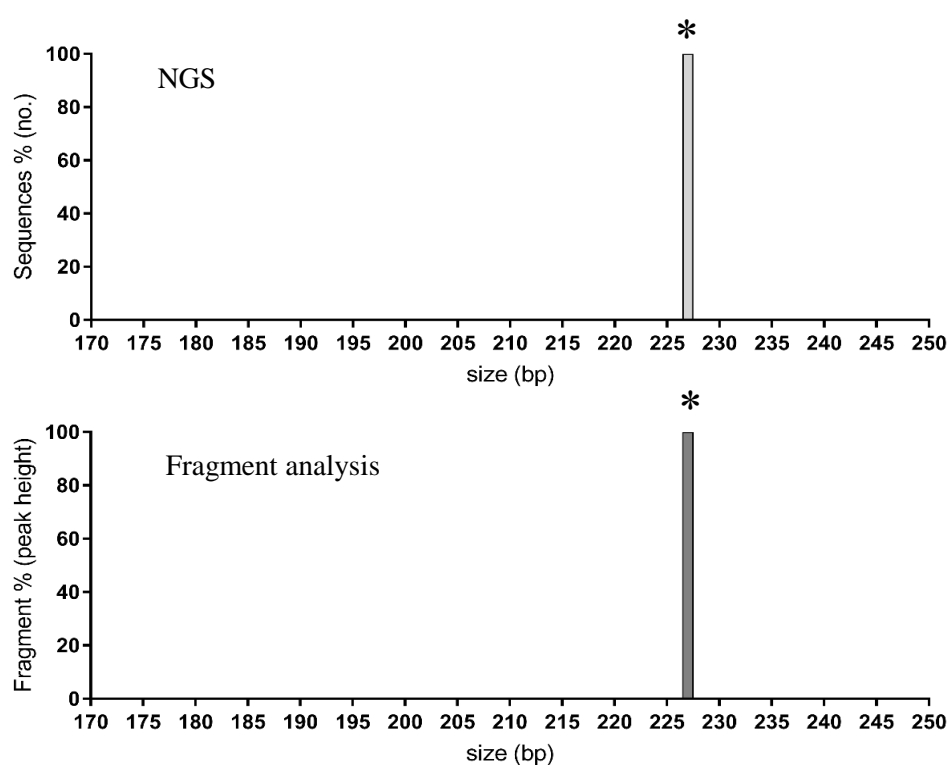
Figure 6.19. Representative results of detected fragments by NGS and fragment analysis. (A) WT control (control No. 7), (B) *piwil2* mutant which does not possess WT length fragment (mutant 11, phenotype A) and (C) *piwil2* mutant which has WT length fragment (mutant 21, phenotype B). WT length fragment is denoted by asterisk (*). (D) the average number of detected fragments in 52 mutants by fragment analysis and NGS. NGS_size: number of sequences with different indel size; NGS_alleles: number of sequences with different alleles. Data are shown as mean ± SEM. Superscripts denote statistically significant difference between groups ($p < 0.05$). (continued on next page)

(B) Mutant without WT length fragment (mutant 11, phenotype A)
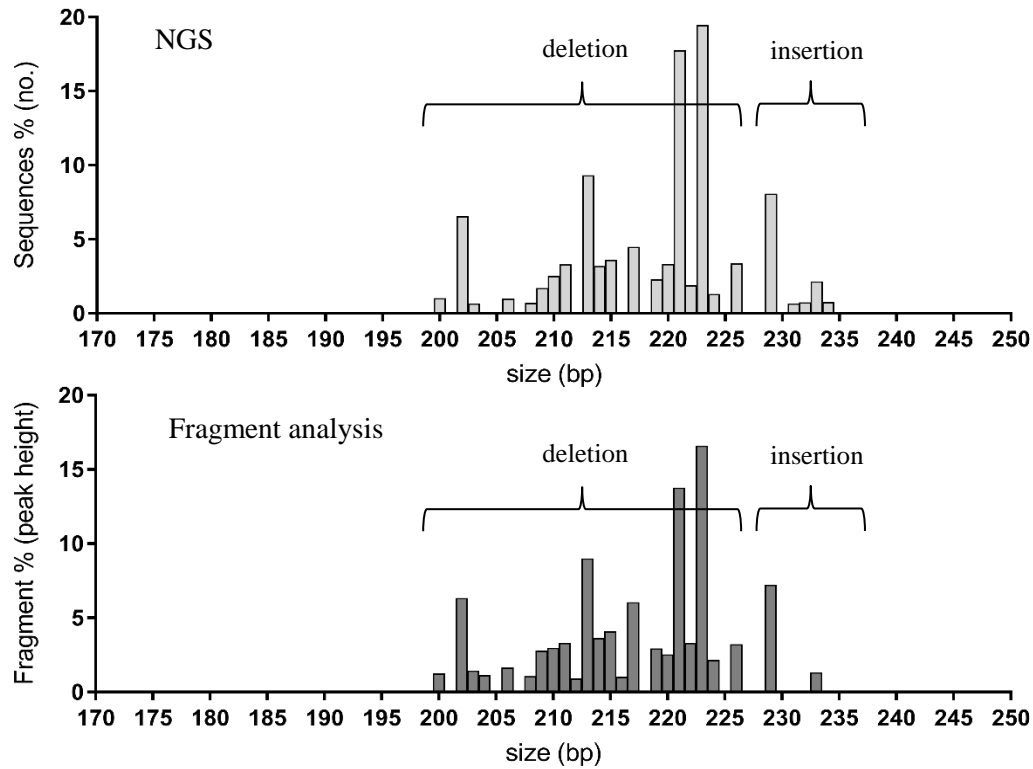
(C) Mutant with WT length fragment (mutant 21, phenotype B)
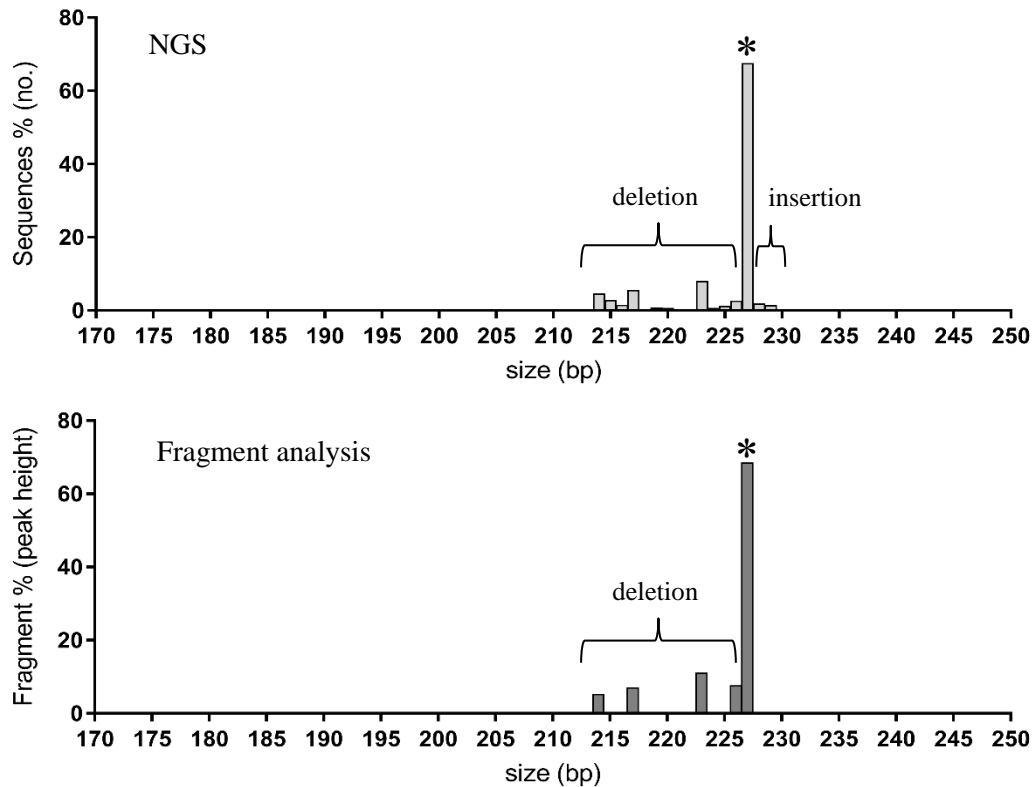
Figure 6.19. (continued on next page)

(D) Average number of detected fragments in 52 mutants by fragment analysis and NGS
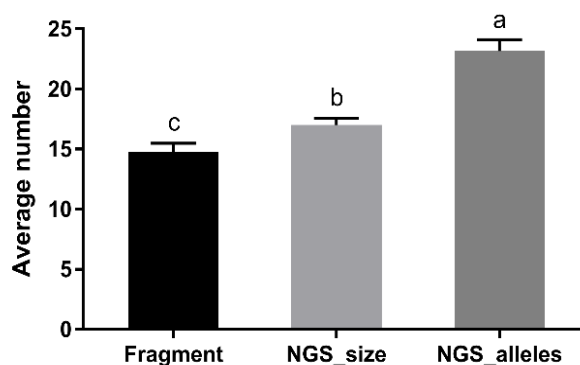
Figure 6.19. (continued from previous page)

6.4 Discussion

This study took a panel of gDNA samples extracted from previously generated KO *piwil2* mutant Nile tilapia and used this to compare the ability of the most frequently used mutation screening methods (T7E1, RGEN, HRMA, fragment analysis and NGS) to characterise the individual mosaic genotype. Thereafter, a secondary aim was to improve our understanding of the possible genotype to phenotype association in CRISPR/Cas9 *piwil2* mutants. At its conclusion, it was apparent that in terms of characterising mosaic genotypes only the fragment and NGS analysis methods were capable of describing genotype diversity to a level that potentially could help to further understand phenotypic effects. However, on further consideration of the outputs of these methods, there was no apparent diagnostic genotype that could clearly explain differences in observed phenotypes (PGC presence) in KO *piwil2* mutant Nile tilapia.

To date, CRISPR/Cas9-mediated sterility studies lacked a comprehensive understanding of genotypes generated in F0 animals due to the biased mutant screening methods, lacking standardisation and methodological details. The most frequently used screening methods in sterility-related gene functional studies were RED and Sanger sequencing (Chen *et al*., 2017; Gay *et al*., 2018; Li *et al*., 2014; Sawamura *et al*., 2017), and a few studies adopted HRMA or SURVEYOR (Qin *et al*., 2016; Zhang *et al*., 2015b). Some studies lack any confirmation of mutant sequences including *dnd* KO salmon (Wargelius *et al*., 2016), *sf-1* KO tilapia (Xie *et al*., 2016) and *fshr* & *lhcgr* double KO zebrafish by TALEN (Zhang *et al*., 2015b). In previous tilapia genome editing studies

using CRISPR/Cas9 targeting gonad-related genes, RED and Sanger sequencing were utilised using pooled F0 embryos (Feng *et al.*, 2015; Jiang *et al.*, 2016, 2017; Li *et al.*, 2015; Xie *et al.*, 2016; Zhang *et al.*, 2014b). However, RED requires restriction enzyme site on the target sites; therefore, it is certainly not applicable for most studies. In addition, pooling of samples and a small scale of sequencing can be inappropriate and insufficient measures. There were no studies reporting the methodological rigour on the selection of such methods for the screening. Therefore, this study was dedicated to compare and validate the different mutant screening methods in individual F0 fish.

First of all, all the methods tested in this study successfully distinguished mutants from WT. T7E1 and RGENs were able to detect mutants, but they were not able to provide details of the genotype diversity. Both assays alluded to a greater level of mutation in phenotype A, *i.e.* higher arbitrary gene modification rates in phenotype A than B and C. However, this is not a reliable diagnostic measure to characterise the phenotypes as the actual mutation rate calculated by NGS was not different between phenotype groups. Importantly, it was found that the mutation frequencies detected by T7E1 were consistently lower than the mutation rate determined by NGS in this study. Significant underestimation of the arbitrary gene modification rate was also reported by Kim *et al.* (2014a), where they compared T7E1 and RGEN assays to assess the accuracy of the estimated mutation rates using known mutation samples made by mixing a mutant clone with WT in various ratios from 0 to 100%. The cleavage of RGENs was proportional to the WT to the mutant ratio, while the T7E1 assay poorly reflected the mutant ratio at levels of inclusion higher than 20% (Kim *et al.*, 2014a). This might be due to inefficient heteroduplex formation and overlooking homoduplex formation between mutant alleles in the calculation of mutation rate using T7E1 (Kim *et al.*, 2009; Renaud *et al.*, 2016). Being based on enzyme reactions, incomplete digestion due to suboptimal reaction conditions can cause false negative results in T7E1 assay or false positive results in RGENs. The mutation rate calculated by RGENs showed a moderate correlation with NGS which is in agreement (Kim *et al.*, 2014a). Unlike T7E1, RGEN is not affected by a high rate of polymorphism in the genome (Kim *et al.*, 2014a). However, it should be acknowledged that false negative results can occur when there is only a small change in the target sequences, preferentially non-seed sequences (seed sequence indicates 12 bp adjacent to 5' of the PAM), which can be re-cut by the sgRNA/Cas9 complex (Cong *et al.*, 2013; Jinek *et al.*, 2012). It was also noted that RGEN reaction required a long incubation time of 16 hrs and still a few of the control samples were not

fully digested, showing a high probability of false positive results. There have been many studies which adopted the enzymatic digestion methods including T7E1, SURVEYOR, RED and RGEN to assess mutation frequencies using pooled embryo samples; however, it is recommended to not pool samples as well as to compare the mutation frequencies calculated by different methods.

HRMA detects the changes of GC content, length and heterozygosity by using saturating dyes and high resolution melt curve program in specific analysis platforms (Reed *et al*., 2007). The method has a high sensitivity which can detect one bp indels (Denbow *et al*., 2017) and has the ability to screen F0 mutants in a high throughput fashion. In the current study mutants were typically seen to possess two to three Tm peaks and based on the melt curve shapes and Tm clusters, the HRMA melt curves were grouped into three types. There was a tendency for phenotype A to display a high proportion of melt curve type I, which has multiple dominant Tm peaks including one in the lowest Tm cluster. The melt temperature and the shape of the melt curves, however, were not able to provide any clear insight into the genotypes. In fact, the information obtained by HRMA was as limited as normal qPCR melt curve analysis used in the previous Chapter 4 & 5 of this thesis. The wide usage of HRMA was reported in various genome editing studies including CRISPR/Cas9, TALEN and ZFN from plants, mammalian cell lines to fish (Chen *et al*., 2018; Dahlem *et al*., 2012; Denbow *et al*., 2017; Samarut *et al*., 2016; Thomas *et al*., 2014; Wang *et al*., 2015; Zhang *et al*., 2015b). In these studies, HRMA was demonstrated to be rapid and highly efficient at detecting indel variants at the targeted loci in F0 animals with high throughput, but no more genotypic information could be obtained. Unlike typical qPCR melt curve analysis, HRMA requires proprietary reagents containing a saturating fluorescent dye which is suitable for the precise collection of fluorescence data in melt curve analysis (Tse *et al*., 2011). However, some studies have demonstrated that SYBR green also can be used to perform an equivalent analysis to HRMA (D'Agostino *et al*., 2016; Price *et al*., 2007). Typical qPCR melt curve analysis or HRMA program with SYBR green dye would be time- and cost-effective methods to detect the heteroduplex as an initial screening for F0 generation, but it should not be considered to be informative beyond this.

Fragment analysis competently detected mutants in F0 animals in high throughput fashion. This analysis had previously been validated in mutant cell lines generated by CRISPR/Cas9, showing that it is able to detect small indel with 1 bp resolution as well as quantify the proportion of mutant fragments (Ramlee *et al*., 2015; Yang *et al*., 2015).

This assay was used in zebrafish F0 to determine target specific efficiency of sgRNA and it showed a strong correlation with germline transmission efficiency (Carrington *et al*., 2015). Contrary to T7E1, RGENs and HRMA, this analysis was not restricted to the detection of mutants and the estimation of mutation rate, but also provided insight into the diversity and composition of various indel sizes in mutant F0 animals. Compared to NGS, 70% of indel variants detected by NGS were matched with fragment analysis and the arbitrary gene modification rate calculated by fragment analysis showed the highest correlation with the mutation rate measured by NGS. The total number of fragments detected by fragment analysis was significantly lower than NGS due to lower sensitivity and the inability to detect mutant sequences that had not changed in physical length (*i.e.* WT length mutated sequences). It is also important to note that the nucleotide composition of the amplicon, the property of dye and the resolution range of the capillary filling matrix can all impact on the movement of the amplicon within the gel capillary, which can then cause incorrect size determination of the fragments as suggested by Mansfield *et al*. (1996). Therefore, in the context of CRISPR/Cas9 studies, the data collected by fragment analysis is not suitable to assess frame-shift mutation rates or determine the total number of different alleles on target site due to the resolution error of this assay. Overall, fragment analysis offers an overview of the indel size variants, but the wide application of this method in mutant screening is hampered by the inability to provide sequence information.

In-depth analysis of the deep sequencing data revealed that there is a high level of mosaicism in F0 animals produced by CRISPR/Cas9 and subsequent NHEJ repair. In 52 F0 samples, the frequency of deletion notably exceeded the frequencies of insertion and substitution. In addition, deletion events occurred over a wider range than insertion or substitution. In accordance with the higher frequency of deletion, the major indel size ranged between -14 and 1 bp (25th to 75th percentile), with a clear bias towards deletion in frequent alleles, as also reported in mice and human cell lines (Mali *et al*., 2013; Parikh *et al*., 2015). This is a notably wider size range than has been reported in previous studies where the most frequent size of indel produced by CRISPR/Cas9 and subsequent NHEJ repair was 1 bp insertion (54.1%) analysed by Sanger sequencing (5 - 6 clones per sample) (Ma *et al*., 2015), 1 bp insertion and 1 – 3 bp deletions (64%) by Sanger sequencing (10 clones per sample) (Pan *et al*., 2016), 1 bp deletion (67%) by NGS (Seeger and Sohn, 2016) or 5 bp deletion (42%) by NGS (Wang *et al*., 2016). The current study is the first to report deep sequencing data in sterility-related gene functional analysis, and it was

noted that small scale Sanger sequencing cannot capture the full spectrum of variants present in F0 animals. Given the level of mosaicism, various indel variants and frequencies, it is recommended to genotype individual samples rather than pooled samples using deep sequencing platform. In addition, the cost of NGS assay can be reduced when it is run along with large multiplexed samples, *e.g.* 3% loading in a full run is *pro rata* costs about £30.

Although NHEJ is prone to introduce small deletions or insertions to the target site (Kim and Kim, 2014), 36.5% of mutants in the present study showed the presence of putative large size deletion (90 bp – 670 bp) and insertion (530 – 580 bp) events when the amplicon spanning 719 bp upstream and 251 bp downstream from the canonical cut site. While the NGS analysis is restricted to a smaller fragment size, it too provided evidence of significant deletion and insertion events ranging from -82 bp to +124 bp in the WT amplicon size of 211 bp. The appearance of large indels (-77 to +96 bp) has previously been reported in CRISPR/Cas9 studies in zebrafish, killifish and Atlantic salmon (Aluru *et al*., 2015; Edvardsen *et al*., 2014; Jao *et al*., 2013), but the incidence of large indels is unknown. On the other hand, two nearby sgRNAs-mediated large deletions of 42.6 kb were reported in mice, where around 20% of mutant mice possess large indels of > 600 bp ($n = 29$) (Parikh *et al*., 2015). Equally, it was reported that NHEJ-mediated knock-in of kilobase-sized donor DNA is also possible (Salsman and Dellaire, 2017). Together, these results indicate that such large-scale deletions or insertions generated by NHEJ repair following DSBs are not negligible in CRISPR/Cas9-mediated genome editing. Such event, if it is scalable, might have a potential use to induce a large deletion in the target site. It is necessary to more comprehensively characterise the extent of such large-scale deletion events and to this end, the presently used NGS sequencing approach would not be appropriate due to the limited sequencing read length. Alternative sequencing methodologies which enable long reads of targeted regions would have to be adopted, for example, the range of methods proposed by Zischewski *et al*. (2017) to detect off-target effects would be relevant in this context. In the first instance, it would be proposed that the most pragmatic approach would be to combine targeted library amplification, as performed herein, of fragments several kb in size with sequencing platforms which give long read lengths. Such platforms include single molecule real-time (SMRT) DNA sequencing (commonly referred to as PacBio), with average read length of 8.5 kb (Hendel *et al*., 2014) or nanopore sequencing platforms, *e.g.* MinIon by Oxford Nanopore technologies, where average read length can be in excess of 10 kb, with the

latter being the most accessible and cost effective (Jain *et al*., 2016). If such methodologies can be validated then it will be possible to genuinely explore the large indel effects of CRISPR/Cas9 gene editing to determine if these effects are reproducible, predictable and in the future could be harnessed as an additional research methodology.

Based on the results from NGS and fragment analysis, it was possible to explore for the first time the possible association between individual genotype diversity and the observed phenotypic differences in *piwil2* KO larvae. The analysis revealed that 26 out of 52 *piwil2* KO larva (50%) showed 100% biallelic mutation rate. The biallelic KO of a target gene in F0 was also demonstrated in medaka and salmon and results showed that 50% of *dnd* KO F0 medaka (2/4) and 20% of *slc45a2* KO F0 salmon (1/5) had 100% biallelic mutation (Edvardsen *et al*., 2014; Sawamura *et al*., 2017). However, the observation was conducted at a very limited scale (*e.g*. Sanger sequencing of 12 – 17 clones per *dnd* KO medaka, $n = 4$ and 50 – 100 clones per *slc45a2* KO salmon ($n = 5$) (Edvardsen *et al*., 2014; Sawamura *et al*., 2017). Therefore, the depth of coverage of the NGS result in the current study, conducted in 52 mutants of *piwil2* with an average of over 10K paired-end reads per fish, significantly enhances the comprehension of the diversity of indels in F0 animals and demonstrates the high efficiency in producing 100% biallelic mutants in F0. The results from the present study showed no significant difference in the average mutation rates between phenotypic groups which contrasts with the findings of Edvardsen *et al*. (2014) and Shigeta *et al*. (2016), who reported that a higher degree of mutation was associated to the severity of the phenotypic response. The potential disparity might be due to the difference in mutation analysis between these studies or the difference in target genes. Given that the previous studies were conducted in a limited number of samples without deep sequencing, the results might be biased or some aspects may have been overlooked. Alternatively, another potential explanation may be differences in the properties of the target area important for the function of the gene. The high mutation rate without significant difference among the different phenotype groups might indicate that the resultant changes in the target sequence are decisive to the phenotypic difference, rather than the simple presence of indels (*i.e.* mutation rate) in the target area of *piwil2* gene. In addition to total mutation rate, the frame-shift mutation rates were investigated as the different proportions of in-frame mutation existing in mutants can generate partial loss-of-function in proteins and impacts on the severity of the phenotype (Jao *et al*., 2013). However, frame-shift mutation rates were not apparently correlated with the phenotypes in this study. No clear correlation

between frame-shift mutation rate and the severity in phenotypes was reported in red sea bream (*Pagrus major*) (Kishimoto *et al*., 2018). Considered that the target area was located in the downstream of *piwil2* (21st exon out of a total of 24 exons), the frame-shift mutation may not make an apparent difference in the phenotype in this case.

While the total mutation and frame-shift mutation rates did not clearly correlate with the observed phenotype groupings, the indel size distribution analysed by both fragment analysis and NGS showed that phenotype A had significantly higher proportion of > 5 bp deletions. Equally, there were significantly higher frequencies of deletion events around the canonical cut site in phenotype A than in B and C. In addition the number of different alleles was significantly higher in phenotype A than B and C. Together, this is suggestive of an association in severity of mutation effects and the degree of phenotypic response in *piwil2* KO larvae. However, specific diagnostic features in the genotypes are not apparent that could be used to reliably predict the phenotypic response in these animals. As discussed in the previous Chapter, a potential cause for the lack of clear associations between genotypes and phenotypic responses may be the phenotyping which had to be performed prior to first-feeding for Home Office reasons. Future studies should aim to assess the phenotypes of the animals after they reach puberty. Equally, it is possible that the genotype as interpreted from the gDNA extracted from the head/tail is not a true reflection of the genotype of the PGC lineage within the animal. Oliver *et al*. (2015) and Brocal *et al*. (2016) demonstrated that in mice and zebrafish, somatic tissues are not predictive of the alleles (or their frequencies) that are present in gamete samples of the same individuals. In addition, Kishimoto *et al*. (2018) reported the different frequencies of each variant among tissues (*e.g.* brain, liver, muscle, gonad and pectoral fin) within individual red sea bream. In the current study, there was a limitation in the analysis of germline genotypes as *piwil2* KO mutants may not possess PGCs which is a technical hurdle that cannot be overcome. However, more broadly, the potential diversity in cell population genotypes is possibly due to the different efficiencies of CRISPR/Cas9 system in different cell types as well as the ontogeny of cell development in the animals resulting in different cell populations with different mosaic genotypes in the same animal. This phenomenon has not been deeply investigated to date; therefore, further research is required to enhance the understanding of the genotypes induced by CRISPR/Cas9.

In summary, although T7E1 and RGENs were easy to operate, they would not be easily applied for large-scale screening and the mutation frequencies calculated by T7E1

were not actual mutation rates due to the inefficient heteroduplex formation, overlooking of homoduplexes for the mutation rate estimation and possible incomplete digestion of nucleases (Bae *et al*., 2014; Kim *et al*., 2009; Renaud *et al*., 2016). HRMA, fragment analysis and NGS can be conducted in a high throughput manner. All methods tested here reliably identified the mutants, but T7E1, RGENs and HRMA were not informative as to the correlation between the complex mosaic genotypes and the severity of the phenotypes. Fragment analysis could capture and map indel size spectrum and the mutation rate which was the closest to the actual mutation rate analysed by NGS among all methods tested. However, the limited resolution of size for detection and the lack of sequence details hinder the clear depiction of the mutation diversity in mutants. Certainly, NGS is the most informative, accurate and high throughput screening method among the five methods tested. As demonstrated in this current study it is possible to access the power of NGS genotyping in a cost-effective manner if low volume sequencing libraries can be multiplexed into other sequencing experiments. It was also noted that longer read length of deep sequencing would be required considered that there are large indels generated by CRISPR/Cas9. Further studies are required to find out the genotypic features which can effectively induce phenotypic changes to enable reliable and precise gene functional analysis in F0 animals created by genome editing.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSION

7.1 Overview

The current study aimed to contribute to the development of a new means of sterilisation in fish. The overall goal of this work was to demonstrate that sterile fish can be produced by targeting PGCs using gene editing. Sterility in fish is a highly desirable trait in aquaculture for sustainability (*e.g.* protection of wild stocks) and productivity (*e.g.* enhanced performances) reasons. Within this context, this thesis consisted of a series of methodological and experimental chapters directed towards the development of a reliable and effective genome editing approach to produce sterile Nile tilapia. The thesis is organised around five original scientific chapters to initially screen for candidate genes during the ontogenic development of Nile tilapia (Chapter 2), study germ cell death using heat treatment (Chapter 3), optimise the CRISPR/Cas9 platform targeting *piwil2* gene (Chapter 4), test the functional effect in the *piwil2* edited mutants (Chapter 5) and finally, study the complex genotypes generated by CRISPR/Cas9 (Chapter 6). The starting point of this thesis was the screening of a panel of 11 candidate genes, selected based on their putative role in PGC survival and maintenance. Their expression profiles were studied during early embryonic development and in various tissues from adult Nile tilapia. Based on results from both Chapters 2 and 3, the *piwil2* gene was chosen as the preferred target for functional studies as it is a novel gene showing a potential role in germ cell survival. Thereafter, in Chapter 4, the genome editing approach targeting *piwil2* using the CRISPR/Cas9 platform was optimised through an improvement of the microinjection system and validation of optimal sgRNA concentration. This study resulted in a method that reliably produced > 90% mutant embryos. Then, in Chapter 5, the physiological impact of *piwil2* KO on PGCs was examined using the optimised genome editing method. Unfortunately, the confirmation of mutant phenotypes could not be done through GFP labelling of PGCs and therefore relied on histological analyses at pre-first feeding stage. While this technique may not be as efficient as GFP labelling, three different mutant phenotypes were clearly distinguished. The next important question was related to linking these phenotypes to the genotypes of *piwil2* KO mutants. Therefore, Chapter 6 was dedicated to the characterisation of the genome editing effects by analysing deep sequencing data to reveal the subsequent individual genotype diversity generated by CRISPR/Cas9. Overall, the current thesis provides new scientific knowledge and supporting evidence for the use of the CRISPR/Cas9 gene editing platform to study gene function associated with sterility and induce sterilisation in Nile tilapia.

7.2 Limitation and future perspectives for gene functional analysis in F0 using CRISPR/Cas system

CRISPR/Cas9 mediated genome editing is highly efficient so that complete KO can be achieved in F0, which is essential in the context of induced sterility. For the direct functional analysis in F0, it is important to have high mutation rates as well as accurate phenotyping in F0. In the present thesis, the optimised CRISPR/Cas9 system induced on-target mutation in *piwil2* in over 95% of injected eggs, of which half had 100% on-target mutation; however, the histological phenotyping showed a variation in the severity of the resulting phenotypes based on PGCs presence and abundance (Chapter 6). It should be acknowledged that the histological observation used in the current study cannot definitively confirm the complete absence of PGCs; therefore, a more reliable phenotyping method for PGCs is required for future gene editing studies such as generating a transgenic line in which PGCs are labelled by fluorescent protein for research purpose (Gross-Thebing *et al*., 2017). In addition, ISH was attempted in the present work but unsuccessful due to the low expression of *vasa* and time constraints.

The in-depth sequencing analysis revealed the complexity of the mosaic genotypes generated by the CRISPR/Cas9 treatment, but unfortunately, this did not provide a clearer insight to aid the interpretation of the observed phenotypic effects (Chapter 6). Despite the many sterility studies using gene editing published recently in fish, none provided in-depth genotyping results from F0 animals and the selected screening methods (*e.g.* RED, PCR amplification) lacked standardisation and methodological details (Chen *et al*., 2017; Gay *et al*., 2018; Li *et al*., 2014; Qin *et al*., 2016; Sawamura *et al*., 2017; Wargelius *et al*., 2016; Xie *et al*., 2016; Zhang *et al*., 2015b). Mianné *et al*. (2017) reviewed a limited selection of screening methods including PCR amplification, gel electrophoresis and Sanger sequencing to find F0 with desired indels or screen subsequent F1 offsprings from edited F0 parents. The authors acknowledged that the complexity of CRISPR/Cas9 mosaic genotypes is underestimated in many studies and advised to perform individual Sanger sequencing. However, the use of Sanger sequencing in that study was limited in its depth of coverage (up to 24 reads per individual) which is typical of the literature in this field. In the current thesis, the comparison of screening methods including T7E1, RGEN, HRMA, fragment analysis and deep sequencing from 52 *piwil2* mutants showed that the most reliable screening method is deep sequencing. This approach provided a

detailed insight into the indel spectrum induced by CRISPR/Cas9, thanks to the comprehensive depth of coverage (*circa* 10K paired-end reads per individual) in the most cost effective manner (*circa* £0.6 per individual *pro rata*) by multiplexing into a large sequencing run. Since it was noted that there are clear limitations in the other screening methods, a careful consideration should be given to the selection of the screening method in the targeted genome editing studies and where possible, the use of high throughput sequencing methods should be adopted.

The in-depth sequencing results and histological observations of PGCs confirmed that even while almost 100% KO larvae can be obtained, the phenotypic results remained variable and unpredictable, implying that some indels may not cause complete loss of function. This clearly indicates that it is still in an early stage of development to evaluate gene function in F0 and further improvements are required to enable the functional analysis in injected animals. First of all, to obtain predictable and homogeneous gene editing results, the efficiency of HDR-mediated repair should be enhanced (Gutschner *et al.*, 2016). Secondly, the completion of the cleavage activity of Cas9/sgRNA complex in 1-cell stage embryo is desired to minimise the number of different mutant alleles which can be improved via delivery of purified Cas9 ribonucleoproteins (Kim *et al.*, 2014b). In addition, it is important to control the frequency of cutting events by Cas9/sgRNA complex to prevent repeated cleavage in the edited site, as it can cause mosaicism (Kim *et al.*, 2014c). These aspects are not only crucial for functional analysis in F0, but for the application of this technique in various areas including gene therapy.

Another important aspect to be considered for effective gene functional analysis is the sgRNA design strategy. A number of studies designed sgRNAs around the start codon, aiming to disrupt the whole ORF; however, it should be noted that approximately one third of the mutated sequences would be in-frame mutations and cells with frame-shift mutations could use alternative start codons resulting in truncated proteins (Prykhozhij *et al.*, 2017). To prevent this, a more effective method would be to employ two sgRNAs to induce large deletions in ORF (Sawamura *et al.*, 2017), but the efficiency of each sgRNA would require validation prior to targeting multiple sites as it can result in large variations as shown in the present study (Chapter 4). In addition, a functional role of specific amino acid sites will be more precisely predicted when highly efficient HDR repair is available for the phenotyping of F0. The improvement in accuracy and reliability of genome editing techniques such as CRISPR/Cas system is an important area for future research.

7.3 Sterilisation of Nile tilapia

One of the major problems in Nile tilapia farming is the frequent breeding in ponds stocked with a mixed sex population which hinders the growth of fish by causing overcrowding and leading to a high proportion of small sized fish at harvest (Hulata et al., 1983). Therefore, farming of monosex projeny is desired. As monosex male tilapia progeny shows less energy diversion into gametogenesis than monosex females, the industry requires the supply of monosex males, which can be obtained by manual sorting, hybridisation, hormonal treatment or YY/GMT (Beardmore *et al*., 2001). The most common method used is hormonal treatment as it is cheap and effective; however, this method is banned in most economically developed regions including the EU (Council of the European Union, 2003; Little and Hulata, 2000). Importantly, hormonal sex reversal does not solve the potential threat caused by escapees. Reproductive containment in farmed fish is highly recommended to ensure sustainability through the prevention of potential cross-breeding between escaped farmed fish and wild stocks and enhance productivity through the suppression of reproduction (Muir and Howard, 1999; Wong and Zohar, 2015a).

Effective and reliable disruption of PGCs could become a new means of inducing sterility in farmed stocks in the future although it is still early days. The main problem in the development of sterile fish is that the "sterile trait" cannot be considered heritable *per se* (Zhang *et al*., 2015a). To circumvent this, there have been several strategies tested to maintain broodstock fertility through transgenic approaches which seek to induce sterility in the F1 progeny through manipulated, but fertile, F0 parents. For example, one transgenic line has a transcriptional activator, GAL4, which binds to UAS (upstream activation sequence) to activate the target gene (Brand and Perrimon, 1993), and the other transgenic line is designed to express antisense *dnd* RNA controlled by UAS so that the antisense *dnd* is activated in the subsequent progeny by the Gal4/UAS system to disrupt PGC migration (Zhang *et al*., 2015a). Another example is to generate a single transgenic line which has a foreign gene construct consisting of a controllable promoter (*e.g. hsp70*) and the target gene, of which overexpression induces sterility. Wong and Collodi (2013) tested a zebrafish transgenic line which expresses *cxcl12a* under control of *hsp70* promoter. High water temperature (34.5°C for 18 hours) induced the overexpression of *cxcl12a* and resulted in the disruption of normal migration of PGCs in all exposed

embryos, which became infertile adults. Alternatively, there have been a number of non-transgenic strategies seeking to sterilise fish directly in the F0 generation. Presslauer *et al*. (2014) tested an autoimmune approach against gonadal proteins such as Gsdf, growth differentiation factor (Gdf9), or lymphocyte antigen 75 (Cd205), by injecting fish with anti-Gsdf, anti-Gdf9 or anti-Cd205, resulting in the inconsistent suppression of gonadal maturation. Interestingly, Wong and Zohar (2015b) reported a bath immersion method for fertilised embryos with a *dnd* morpholino oligomer (MO) which is carried by a molecular transporter called Vivo, showing high sterility in treated embryos although increased mortalities may be a concern. Overall, these findings suggest sterility could be controlled as a trait in broodstock management.

In both transgenic and non-transgenic approaches, the common bottleneck is the finding of one or several suitable target gene(s) or protein(s) to be disrupted. As it was discussed earlier, *cxcl12a* or *dnd* were reported to be involved in various processes (Ratajczak *et al*., 2006; Škugor *et al*., 2014b). Therefore, it is crucial to find the optimal target genes as well as germline specific regulatory regions for spatial control of the expression *e.g. nanos3* or *vasa* 3'UTR. As shown in the present thesis, a potential target gene such as *piwil2* can be evaluated by the targeted genome editing technique and the knowledge obtained from gene functional analysis can prove to be critical in the development of both transgenic and non-transgenic sterilisation methods. However, further confirmation of sterility in adult fish and a thorough assessment of the potential pleiotropic effects of *piwil2* KO are required prior to any potential application.

Further research is also required to improve the practicability of the technique by addressing production bottlenecks (*e.g.* cost-effective and efficient delivery methods of biomolecules to the gametes/embryos) and upscaling of the induction protocol (Kumari *et al*., 2017; Wong and Zohar, 2015b). In particular, the development of efficient and reliable delivery methods is essential for commercial application as well as research purposes, as microinjection is highly labour-intensive. One of the possible strategies would be sperm transfection of CRISPR constructs as tested in chicken (Cooper *et al*., 2018) or bath immersion of eggs or embryos as demonstrated in zebrafish (Wong and Zohar, 2015b). However, tilapia being a mouthbrooder, this technique would require a change in commercial practices as most farms collect larvae from the female's mouth or fry in the breeding ponds. They would need instead to strip broodstock and perform artificial fertilisation. Results obtained recently including the present work suggest a promising avenue for inducing sterility in tilapia being through transgenic or alternative

non-transgenic treatments. Importantly, this is relevant to most commercially important fish species in which escapee and sexual reproduction remain a problem.

7.4 Potential applications of gene editing techniques in aquaculture

The utilisation of genome editing in combination with genomic selection within modern breeding programs has the potential to rapidly and efficiently improve the genetic traits in farmed animal (Abdelrahman *et al*., 2017; Houston, 2017). As the functional analysis of *piwil2* via CRISPR/Cas9 was performed in this study, this gene editing tool can be used to assess causative genetic variances, although it requires significant progress as discussed above to perform a precise and efficient gene editing. These new techniques can then be also used to introduce specific genetic variance related to traits of interest (*e.g.* higher disease resistance, growth performance or flesh quality) into broodstock genome without introducing transgenes (Houston, 2017; Zhu and Ge, 2018). One such example is the porcine reproductive and respiratory syndrome (PRRS) virus resistant pigs for which the domain responsible for the interaction with the PRRS virus, CD163 SRCR5 (cluster of differentiation 163 the fifth scavenger receptor cysteine-rich domain), can be deleted by genome editing (Burkard *et al*., 2017). In addition, beyond the gene editing application, CRISPR/Cas system can be used for various purposes such as regulation of gene expression or editing of epigenome by using catalytically inactive dead Cas9 (dCas9) (Qi *et al*., 2013; Adli, 2018).

Currently, organisms produced by targeted genome editing techniques such as ZFN, TALEN and CRISPR/Cas are considered as GMOs by the European Union and GM foods or feeds should pass a rigorous safety assessment of the European Food Safety Authority (EFSA) to be authorised in the EU (European Parliament and Council of the European Union, 2003). While the use of genetic engineering in domestic animals requires regulatory approval, more and more research on the various type of Cas effectors and the optimisation for the precise gene editing will address the safety issues related to gene editing. For those animals edited by targeted genome editing techniques which do not possess transgenes and the genetic changes are distinctly smaller than the traditional GMOs, a new regulation will be required (Eriksson *et al*., 2018).

7.5 General conclusion

The overall aim of this thesis was to investigate genes involved in the survival of germ cells and subsequently conduct the functional analysis of candidate genes to ultimately provide the basis for the development of a novel sterilisation technique. Although a reliable means of sterilisation could not be developed in the timeframe of this thesis, the potency of *piwil2* gene KO for induction of sterility was evaluated by CRISPR/Cas9 gene editing technique. In the course of the functional analysis, advances were made for the reliable microinjection of Nile tilapia embryos as a delivery method for biomolecules and it showed no hazardous effect on the survival of the injected embryos. While this microinjection system can be used for gene editing studies in the laboratory, development of a new delivery method for commercial application is required. The demonstration of putative sterility in *piwil2* KO fish suggests that this gene plays an essential role in the survival of PGCs and has a potential to be used for sterilisation in fish. In addition, novel comparative data on widely used mutation screening methods clearly indicate the need for in-depth sequencing. Deep sequencing of 52 *piwil2* mutants showed the highly mosaic genotypes in F0, suggesting that further improvement for predictable genome editing is required for a reliable gene functional analysis in F0. In summary, this study provides new insight into target genes essential for PGCs in fish, and methodological requirements for gene editing studies, which contribute to the ultimate aim of developing an alternative sterilisation method to improve productivity and sustainability in aquaculture.

APPENDIX

Table S1. Recipe of SSTNE (Spermidine-Spermine-Tris-NaCl-EGTA) buffer (1 litre)

Weight or volume	Chemical	Note
17.5 g	NaCl	
6.05 g	Tris Base	
1 mL	EDTA 0.2 M	
76 mg	EGTA	E3889, Sigma Aldrich
72 mg	spermidine	S0266, Sigma Aldrich
52 mg	spermine	S1141, Sigma Aldrich
To 1 litre	MilliQ water	

Autoclave and store at 4°C

Note 1: pH is 9.5 -10.0 (Do not require pH adjustment)

Note 2: Do not vortex at any stage of preparation

Table S2. Staining procedure with Haematoxylin and Eosin and mounting

Step	Solution	Time
1	Xylene (dewax 1)	3 min
2	Xylene (dewax 2)	2 min
3	Absolute Alcohol I	2 min
4	Methylated spirit	1 min 30 sec
5	Wash in tap water	30 sec – 1 min
6	Haematoxylin 'Z'	5 min
7	Wash in tap water	30 sec – 1 min
8	1% Acid alcohol[a]	3 quick dips
9	Wash in tap water	30 sec – 1 min
10	Scott's tap water substitute	1 min
11	Wash in tap water	30 sec – 1 min
12	Eosin	5 min
13	Wash in tap water	30 sec
14	Methylated spirit	30 sec
15	Absolute alcohol II	2 min
16	Absolute alcohol III	1 min 30 sec
17	Xylene (clearing)	5 min
18	Xylene (coverslip)*	Until subjected to mounting & coverslip
19	Pertex mounting	Mounted and covered slides should be left to dry in a fume hood for at least 10-15 mins prior to microscopy

* If there are bubbles on slides, agitate in Absolute alcohol III, then in Xylene (clearing) by up & down.

[a], Recipe of 1% Acid alcohol:

Solution	Volume	Note
Methylated spirit	1000ml	Measure the methylated spirits into a winchester/bottle
Hydrochloric acid	10ml	Carefully add the hydrochloric acid
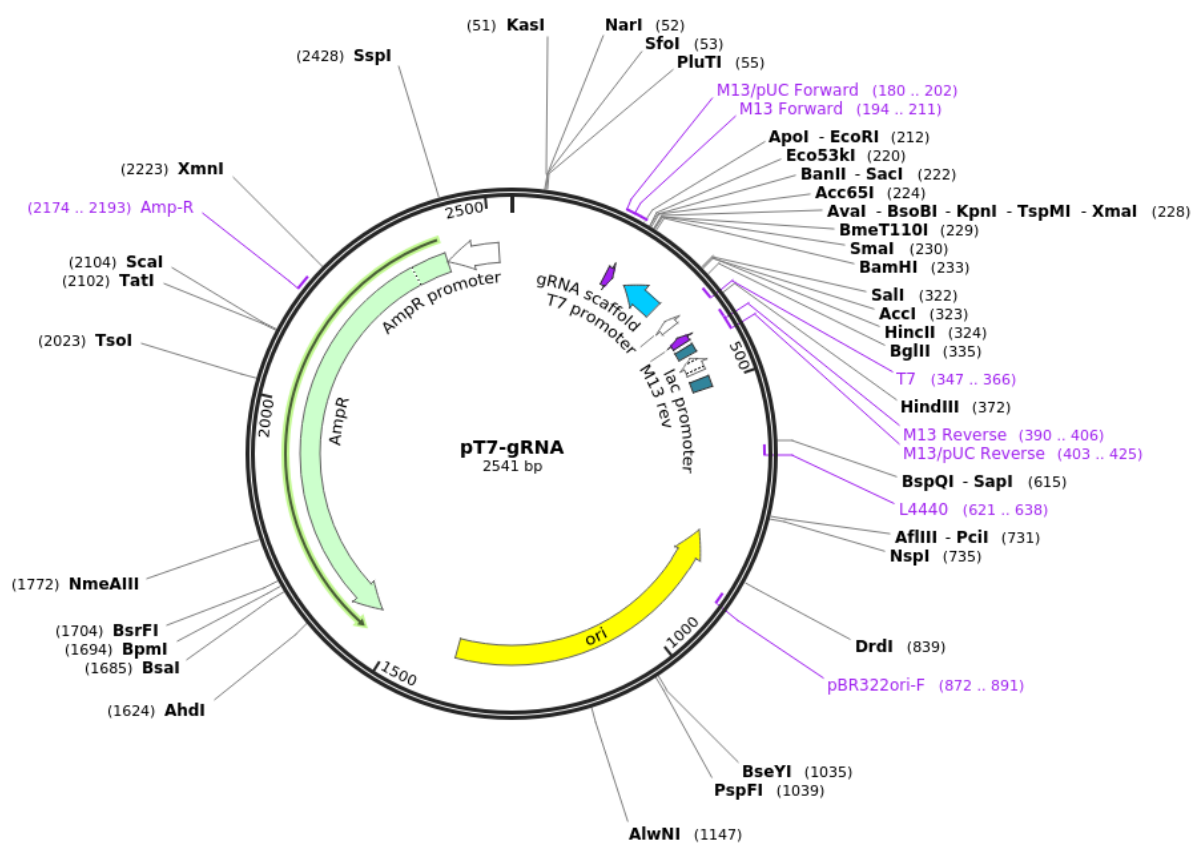
Figure S1. Sequence maps of plasmids used for sgRNA and Cas9 RNA production. (A) pT7-gRNA (Addgene plasmid #46759) and (B) pT3TS-nCas9n (Addgene plasmid #46757) plasmids. Figures frome SnapGene®. (continued on next page)

B
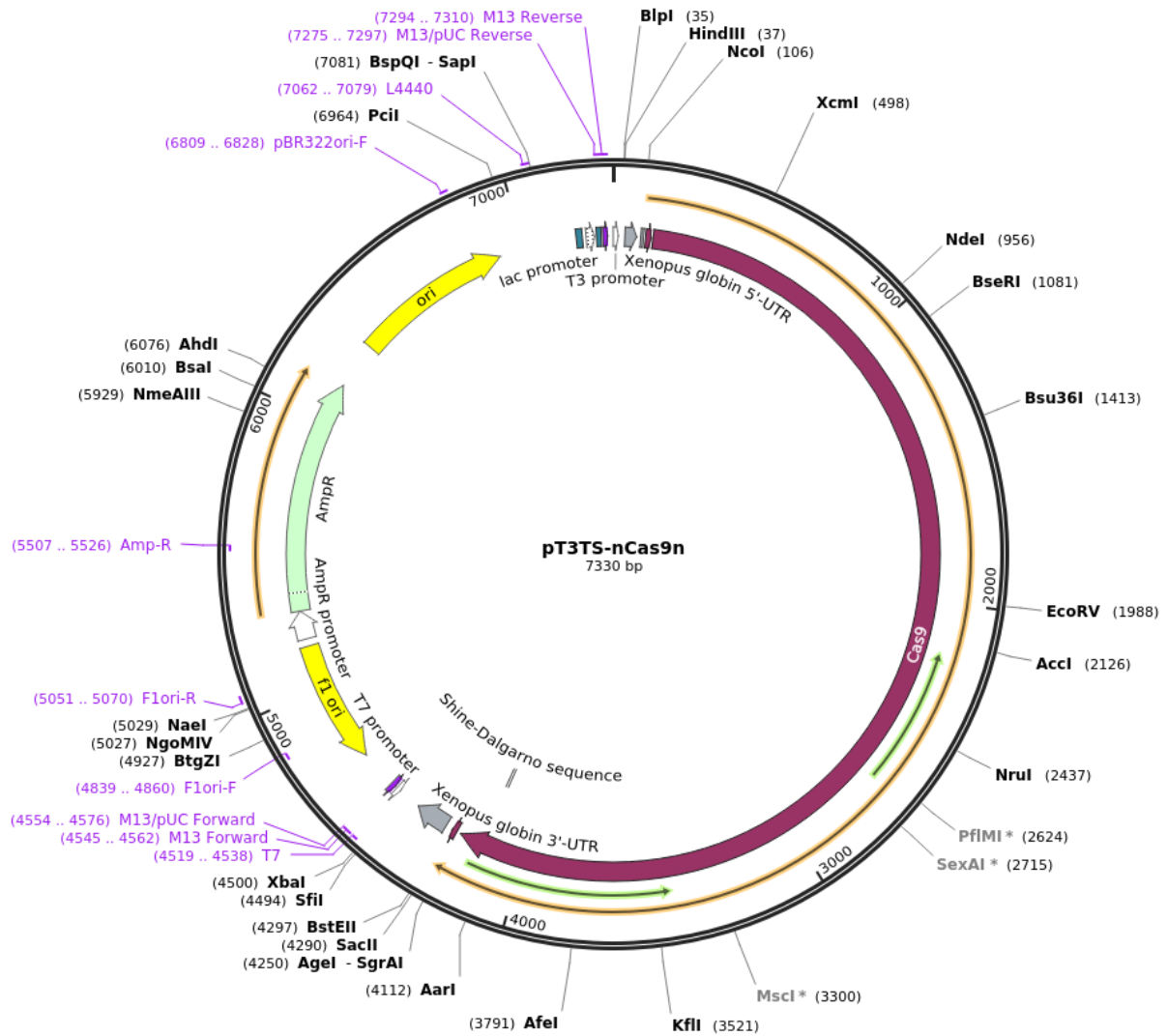
Figure S1. (continued from previous page)

REFERENCES

Aas, G.H., Refstie, T., Gjerde, B., 1991. Evaluation of milt quality of Atlantic salmon. Aquaculture 95, 125–132.

Abdalla, F.C., Cruz-Landim, C., 2004. Occurrence and ultrastructural characterization of "nuage" during oogenesis and early spermatogenesis of *Piaractus mesopotamicus* Holmberg, 1887 (Teleostei). Braz. J. Biol. 64, 555–561.

Abdelrahman, H., ElHady, M., Alcivar-Warren, A., Allen, S., Al-Tobasei, R., Bao, L., Beck, B., Blackburn, H., Bosworth, B., Buchanan, J., Chappell, J., Daniels, W., Dong, S., Dunham, R., Durland, E., Elaswad, A., Gomez-Chiarri, M., Gosh, K., Guo, X., Hackett, P., Hanson, T., Hedgecock, D., Howard, T., Holland, L., Jackson, M., Jin, Y., Khalil, K., Kocher, T., Leeds, T., Li, N., Lindsey, L., Liu, S., Liu, Z., Martin, K., Novriadi, R., Odin, R., Palti, Y., Peatman, E., Proestou, D., Qin, G., Reading, B., Rexroad, C., Roberts, S., Salem, M., Severin, A., Shi, H., Shoemaker, C., Stiles, S., Tan, S., Tang, K.F.J., Thongda, W., Tiersch, T., Tomasso, J., Prabowo, W.T., Vallejo, R., van der Steen, H., Vo, K., Waldbieser, G., Wang, H., Wang, X., Xiang, J., Yang, Y., Yant, R., Yuan, Z., Zeng, Q., Zhou, T., 2017. Aquaculture genomics, genetics and breeding in the United States: current status, challenges, and priorities for future research. BMC Genomics 18, 191.

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E.S., Koonin, E. V, Zhang, F., 2016. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353, aaf5573.

Adam-Guillermin, C., Pereira, S., Della-Vedova, C., Hinton, T., Garnier-Laplace, J., 2012. Genotoxic and Reprotoxic Effects of Tritium and External Gamma Irradiation on Aquatic Animals, in: Reviews of Environmental Contamination and Toxicology. pp. 67–103.

Adams, I.R., McLaren, A., 2002. Sexually dimorphic development of mouse primordial germ cells: switching from oogenesis to spermatogenesis. Development 129, 1155–1164.

Adli, M., 2018. The CRISPR tool kit for genome editing and beyond. Nat. Commun. 9, 1911.

Agarwal, A., Said, T.M., 2005. Oxidative stress, DNA damage and apoptosis in male infertility: a clinical approach. BJU Int. 95, 503–507.

Ain, Q.U., Chung, J.Y., Kim, Y.-H., 2015. Current and future delivery systems for engineered nucleases: ZFN, TALEN and RGEN. J. Control. Release 205, 120–127.

Akingbemi, B.T., Ge, R., Klinefelter, G.R., Zirkin, B.R., Hardy, M.P., 2004. Phthalate-induced Leydig cell hyperplasia is associated with multiple endocrine disturbances. Proc. Natl. Acad. Sci. 101, 775–780.

Ali, P.K.M.M., Rao, G.P.S., 1989. Growth improvement in carp, *Cyprinus carpio* (Linnaeus), sterilized with 17α-methyltestosterone. Aquaculture 76, 157–167.

Aluru, N., Karchner, S.I., Franks, D.G., Nacci, D., Champlin, D., Hahn, M.E., 2015. Targeted mutagenesis of aryl hydrocarbon receptor 2a and 2b genes in Atlantic killifish (*Fundulus heteroclitus*). Aquat. Toxicol. 158, 192–201.

Andersen, C.L., Jensen, J.L., Ørntoft, T.F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245–5250.

Ansai, S., Kinoshita, M., 2014. Targeted mutagenesis using CRISPR/Cas system in medaka. Biol. Open 3, 362–371.

AnvariFar, H., Amirkolaie, A.K., Miandare, H.K., Ouraji, H., Jalali, M.A., Üçüncü, S.İ., 2017. Apoptosis in fish: environmental factors and programmed cell death. Cell Tissue Res. 368, 425–439.

Aoki, Y., Nakamura, S., Ishikawa, Y., Tanaka, M., 2009. Expression and Syntenic Analyses of Four *nanos* Genes in Medaka. Zoolog. Sci. 26, 112–118.

Arai, K., 2001. Genetic improvement of aquaculture finfish species by chromosome manipulation techniques in Japan, in: Reproductive Biotechnology in Finfish Aquaculture. Elsevier, pp. 205–228.

Aravin, A.A., Hannon, G.J., Brennecke, J., 2007. The Piwi-piRNA Pathway Provides an Adaptive Defense in the Transposon Arms Race. Science 318, 761–764.

Asaoka-Taguchi, M., Yamada, M., Nakamura, A., Hanyu, K., Kobayashi, S., 1999. Maternal Pumilio acts together with Nanos in germline development in *Drosophila* embryos. Nat. Cell Biol. 1, 431–437.

Azaza, M.S., Dhraïef, M.N., Kraïem, M.M., 2008. Effects of water temperature on growth and sex ratio of juvenile Nile tilapia *Oreochromis niloticus* (Linnaeus) reared in geothermal waters in southern Tunisia. J. Therm. Biol. 33, 98–105.

Bachvarova, R.F., Crother, B.I., Manova, K., Chatfield, J., Shoemaker, C.M., Crews, D.P., Johnson, A.D., 2009. Expression of *Dazl* and *Vasa* in turtle embryos and ovaries: evidence for inductive specification of germ cells. Evol. Dev. 11, 525–534.

Bae, S., Park, J., Kim, J.-S., 2014. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473–1475.

Baroiller, J.F., D'Cotta, H., Saillant, E., 2009. Environmental effects on fish sex determination and differentiation. Sex Dev. 3, 118–135.

Beardmore, J.., Mair, G.., Lewis, R.., 2001. Monosex male production in finfish as exemplified by tilapia: applications, problems, and prospects. Aquaculture 197, 283–301.

Beer, R.L., Draper, B.W., 2013. *nanos3* maintains germline stem cells and expression of the conserved germline stem cell gene *nanos2* in the zebrafish ovary. Dev. Biol. 374, 308–318.

Bellaiche, J., Lareyre, J.-J., Cauty, C., Yano, A., Allemand, I., Le Gac, F., 2014. Spermatogonial stem cell quest: *nanos2*, marker of a subpopulation of undifferentiated A spermatogonia in trout testis. Biol. Reprod. 90(4), 1–14.

Benfey, T.J., 2016. Effectiveness of triploidy as a management tool for reproductive containment of farmed fish: Atlantic salmon (*Salmo salar*) as a case study. Rev. Aquac. 8, 264–282.

Bhat, N., Hong, Y., 2014. Cloning and expression of *boule* and *dazl* in the Nile tilapia (*Oreochromis niloticus*). Gene 540, 140–145.

Bharadwaj, R., Sharma, L.L., 2000. Effect of methyltestosterone (tablets) in sterilization and masculinization of common carp, *Cyprinus carpio communis* (L.). Indian J. Fish. 47(4), 377–381.

Bharadwaj, R., Sharma, L.L., 2002. Experimental manipulation of sex in *Cyprinus carpio var. communis* by oral administration of testosterone undecanoate. J. Aqua. Trop. 17(1), 7–13.

Billard, R., 1978. Changes in structure and fertilizing ability of marine and freshwater fish spermatozoa diluted in media of various salinities. Aquaculture 14, 187–198.

Billig, H., Chun, S.-Y., Eisenhauer, K., Hsueh, A.J.W., 1996. Gonadal cell apoptosis: hormone-regulated cell demise. Hum. Reprod. Update 2, 103–117.

Blanco-Vives, B., Vera, L.M., Ramos, J., Bayarri, M.J., Mañanós, E., Sánchez-Vázquez, F.J., 2011. Exposure of larvae to daily thermocycles affects gonad development, sex ratio, and sexual steroids in *Solea senegalensis*, kaup. J. Exp. Zool. Part A Ecol. Genet. Physiol. 315A, 162–169.

Blaser, H., Eisenbeiss, S., Neumann, M., Reichman-Fried, M., Thisse, B., Thisse, C., Raz, E., 2005. Transition from non-motile behaviour to directed migration during early PGC development in zebrafish. J. Cell Sci. 118, 4027–4038.

Blaser, H., Reichman-Fried, M., Castanon, I., Dumstrei, K., Marlow, F.L., Kawakami, K., Solnica-Krezel, L., Heisenberg, C.-P., Raz, E., 2006. Migration of zebrafish primordial germ cells: a role for myosin contraction and cytoplasmic flow. Dev. Cell 11, 613–627.

Boekelheide, K., Fleming, S.L., Johnson, K.J., Patel, S.R., Schoenfeld, H.A., 2000. Role of Sertoli Cells in Injury-Associated Testicular Germ Cell Apoptosis. Proc. Soc. Exp. Biol. Med. 225, 105–115.

Bogdanove, A.J., Voytas, D.F., 2011. TAL Effectors: Customizable Proteins for DNA Targeting. Science 333, 1843–1846.

Böhm, I., Schild, H., 2003. Apoptosis: the complex scenario for a silent cell death. Mol. Imaging Biol. 5, 2–14.

Boldajipour, B., Mahabaleshwar, H., Kardash, E., Reichman-Fried, M., Blaser, H., Minina, S., Wilson, D., Xu, Q., Raz, E., 2008. Control of chemokine-guided cell migration by ligand sequestration. Cell 132, 463–473.

Bolger, A. M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

Boonanuntanasarn, S., Bunlipatanon, P., Ichida, K., Yoohat, K., Mengyu, O., Detsathit, S., Yazawa, R., Yoshizaki, G., 2016. Characterization of a *vasa* homolog in the brown-marbled grouper (*Epinephelus fuscoguttatus*) and its expression in gonad and germ cells during larval development. Fish Physiol. Biochem. 42, 1621–1636.

Borg, B., 1994. Androgens in teleost fishes. Comp. Biochem. Physiol. Part C Pharmacol. Toxicol. Endocrinol. 109, 219–245.

Boutin-Ganache, I., Raposo, M., Raymond, M., Deschepper, C.F., 2001. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. Biotechniques 31, 24–6, 28.

Braat, A.K., Speksnijder, J.E., Zivkovic, D., 1999. Germ line development in fishes. Int. J. Dev. Biol. 43, 745–760.

Braat, A.K., van de Water, S., Korving, J., Zivkovic, D., 2001. A zebrafish vasa morphant abolishes Vasa protein but does not affect the establishment of the germline. genesis 30, 183–185.

Brand, A.H., Perrimon, N., 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118, 401–415.

Brawand, D., Wagner, C.E., Li, Y.I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A.Y., Lim, Z.W., Bezault, E., Turner-Maier, J., Johnson, J., Alcazar, R., Noh, H.J., Russell, P., Aken, B., Alföldi, J., Amemiya, C., Azzouzi, N., Baroiller, J.-F., Barloy-Hubler, F., Berlin, A., Bloomquist, R., Carleton, K.L., Conte, M. a., D'Cotta, H., Eshel, O., Gaffney, L., Galibert, F., Gante, H.F., Gnerre, S., Greuter, L., Guyon, R., Haddad, N.S., Haerty, W., Harris, R.M., Hofmann, H. a., Hourlier, T., Hulata, G., Jaffe, D.B., Lara, M., Lee, A.P., MacCallum, I., Mwaiko, S., Nikaido, M., Nishihara, H., Ozouf-Costaz, C., Penman, D.J., Przybylski, D., Rakotomanga, M., Renn, S.C.P., Ribeiro, F.J., Ron, M., Salzburger, W., Sanchez-Pulido, L., Santos, M.E., Searle, S., Sharpe, T., Swofford, R., Tan, F.J., Williams, L., Young, S., Yin, S., Okada, N., Kocher, T.D., Miska, E. a., Lander, E.S., Venkatesh, B., Fernald, R.D., Meyer, A., Ponting, C.P., Streelman, J.T., Lindblad-Toh, K., Seehausen, O., Di Palma, F., 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature. 513, 375–381.

Brem, G., Brenig, B., Horstgenschwark, G., Winnacker, E., 1988. Gene transfer in tilapia (*Oreochromis niloticus*). Aquaculture 68, 209–219.

Brett, J.R., 1979. Environmental Factors and Growth. Fish Physiol. 8, 599–675.

Brinkman, E.K., Chen, T., Amendola, M., van Steensel, B., 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. 42, e168.

Brockhaus, F., Brüne, B., 1999. *p53* accumulation in apoptotic macrophages is an energy demanding process that precedes cytochrome c release in response to nitric oxide. Oncogene 18, 6403–6410.

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E. V, van der Oost, J., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964.

Burkard, C., Lillico, S.G., Reid, E., Jackson, B., Mileham, A.J., Ait-Ali, T., Whitelaw, C.B.A., Archibald, A.L., 2017. Precision engineering for PRRSV resistance in pigs: Macrophages from genome edited pigs lacking CD163 SRCR5 domain are fully resistant to both PRRSV genotypes while maintaining biological function. PLOS Pathog. 13, e1006206.

Burness, G., Moyes, C.D., Montgomerie, R., 2005. Motility, ATP levels and metabolic enzyme activity of sperm from bluegill (*Lepomis macrochirus*). Comp. Biochem. Physiol. Part A Mol. Integr. Physiol. 140, 11–17.

Bushati, N., Stark, A., Brennecke, J., Cohen, S.M., 2008. Temporal Reciprocity of miRNAs and Their Targets during the Maternal-to-Zygotic Transition in *Drosophila*, Current Biology 18 (7), 501–506.

Bye, V.J., Lincoln, R.F., 1986. Commercial methods for the control of sexual maturation in rainbow trout (*Salmo gairdneri* R.). Aquaculture 57, 299–309.

Campolo, F., Gori, M., Favaro, R., Nicolis, S., Pellegrini, M., Botti, F., Rossi, P., Jannini, E.A., Dolci, S., 2013. Essential Role of Sox2 for the Establishment and Maintenance of the Germ Cell Line. Stem Cells 31, 1408–1421.

Carmell, M.A., Girard, A., van de Kant, H.J.G., Bourc'his, D., Bestor, T.H., de Rooij, D.G., Hannon, G.J., 2007. MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. Dev. Cell 12, 503–514.

Carrington, B., Varshney, G.K., Burgess, S.M., Sood, R., 2015. CRISPR-STAT: an easy and reliable PCR-based method to evaluate target-specific sgRNA activity. Nucleic Acids Res. 43, e157.

Chaboissier, M.-C., Kobayashi, A., Vidal, V.I.P., Lützkendorf, S., van de Kant, H.J.G., Wegner, M., de Rooij, D.G., Behringer, R.R., Schedl, A., 2004. Functional analysis of *Sox8* and *Sox9* during sex determination in the mouse. Development 131, 1891–1901.

Chakrapani, V., Patra, S.K., Panda, R.P., Rasal, K.D., Jayasankar, P., Barman, H.K., 2016. Establishing targeted carp TLR22 gene disruption via homologous recombination using CRISPR/Cas9. Dev. Comp. Immunol. 61, 242–247

Chang, C.-F., Lan, S.-C., Chou, H.-Y., 1995. Gonadal histology and plasma sex steroids during sex differentiation in grey mullet, *Mugil cephalus*. J. Exp. Zool. 272, 395–406.

Chabbi, A., Ganesh, C.B., 2017. Influence of cortisol along the pituitary-ovary axis in the cichlid fish *Oreochromis mossambicus*. J. Appl. Ichthyol. 33, 1146–1152.

Chassot, A.A., Le Rolle, M., Jourden, M., Taketo, M.M., Ghyselinck, N.B., Chaboissier, M.C., 2017. Constitutive WNT/CTNNB1 activation triggers spermatogonial stem cell proliferation and germ cell depletion. Dev. Biol. 426, 17–27.

Chateigner-Boutin, A.-L., Small, I., 2007. A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Res. 35, e114.

Chauhan, D., Pandey, P., Ogata, A., Teoh, G., Krett, N., Halgren, R., Rosen, S., Kufe, D., Kharbanda, S., Anderson, K., 1997. Cytochrome c-dependent and -independent induction of apoptosis in multiple myeloma cells. J. Biol. Chem. 272, 29995–29997.

Chen, J., Jiang, D., Tan, D., Fan, Z., Wei, Y., Li, M., Wang, D., 2017. Heterozygous mutation of *eEF1A1b* resulted in spermatogenesis arrest and infertility in male tilapia, Oreochromis niloticus. Sci. Rep. 7, 43733.

Chen, L., Li, W., Katin-Grazzini, L., Ding, J., Gu, X., Li, Y., Gu, T., Wang, R., Lin, X., Deng, Z., McAvoy, R.J., Gmitter, F.G., Deng, Z., Zhao, Y., Li, Y., 2018. A method for the production and expedient screening of CRISPR/Cas9-mediated non-transgenic mutant plants. Hortic. Res. 5, 13.

Ching, B., Chen, X.L., Yong, J.H.A., Wilson, J.M., Hiong, K.C., Sim, E.W.L., Wong, W.P., Lam, S.H., Chew, S.F., Ip, Y.K., 2013. Increases in apoptosis, caspase activity and expression of *p53* and *bax*, and the transition between two types of mitochondrion-rich cells, in the gills of the climbing perch, *Anabas testudineus*, during a progressive acclimation from freshwater to seawater. Front. Physiol. 4, 135.

Chowdhury, I., Joy, K.., 2001. Seminal vesicle and testis secretions in *Heteropneustes fossilis* (Bloch): composition and effects on sperm motility and fertilization. Aquaculture 193, 355–371.

Chuma, S., Nakano, T., 2013. piRNA and spermatogenesis in mice. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 20110338.

Cinalli, R.M., Rangan, P., Lehmann, R., 2008. Germ cells are forever. Cell 132, 559–562.

Cioffi, L.C., Kopchick, J.J., Chen, H.Y., 1994. Production of Transgenic Poultry and Fish, in: Pinkert, C.A. (Ed.), Transgenic Animal Technology: A Laboratory Handbook. Academic Press Inc., San Diego, pp. 297–307.

Cnaani, A., Lee, B.-Y., Zilberman, N., Ozouf-Costaz, C., Hulata, G., Ron, M., D'Hont, A., Baroiller, J.-F., D'Cotta, H., Penman, D.J., Tomasino, E., Coutanceau, J.-P.,

Pepey, E., Shirak, A., Kocher, T.D., 2008. Genetics of sex determination in tilapiine species. Sex Dev. 2, 43–54.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.

Conte, M.A., Gammerdinger, W.J., Bartie, K.L., Penman, D.J., Kocher, T.D., 2017. A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. BMC Genomics 18, 341.

Cooper, C.A., Doran, T.J., Challagulla, A., Tizard, M.L. V, Jenkins, K.A., 2018. Innovative approaches to genome editing in avian species. J. Anim. Sci. Biotechnol. 9, 15.

Cory, S., Adams, J.M., 2002. The bcl2 family: regulators of the cellular life-or-death switch. Nat. Rev. Cancer 2, 647–656.

Council of the European Union, 2003. Directive 2003/74/EC of the European parliament and of the council. Off. J. Eur. Union L 262, 17–21.

Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., Zhang, F., 2017. RNA editing with CRISPR-Cas13. Science 358, 1019–1027.

Cox, D.N., Chao, A., Lin, H., 2000. *piwi* encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development 127, 503–514.

Curado, S., Stainier, D.Y.R., Anderson, R.M., 2008. Nitroreductase-mediated cell/tissue ablation in zebrafish: a spatially and temporally controlled ablation method with applications in developmental and regeneration studies. Nat. Protoc. 3, 948–954.

D'Agostino, Y., Locascio, A., Ristoratore, F., Sordino, P., Spagnuolo, A., Borra, M., D'Aniello, S., 2016. A Rapid and Cheap Methodology for CRISPR/Cas9 Zebrafish Mutant Screening. Mol. Biotechnol. 58, 73–78.

Dahlem, T.J., Hoshijima, K., Jurynec, M.J., Gunther, D., Starker, C.G., Locke, A.S., Weis, A.M., Voytas, D.F., Grunwald, D.J., 2012. Simple Methods for Generating and Detecting Locus-Specific Mutations Induced with TALENs in the Zebrafish Genome. PLoS Genet. 8, e1002861.

Danial, N.N., Korsmeyer, S.J., 2004. Cell death: critical control points. Cell 116, 205–219.

David, N.B., Sapède, D., Saint-Etienne, L., Thisse, C., Thisse, B., Dambly-Chaudière, C., Rosa, F.M., Ghysen, A., 2002. Molecular basis of cell migration in the fish

lateral line: role of the chemokine receptor CXCR4 and of its ligand, SDF1. Proc. Natl. Acad. Sci. 99, 16297–16302.

de Alvarenga, E.R., de França, L.R., 2009. Effects of different temperatures on testis structure and function, with emphasis on somatic cells, in sexually mature Nile tilapias (*Oreochromis niloticus*). Biol. Reprod. 80, 537–544.

de Mateo, S., Sassone-Corsi, P., 2014. Regulation of spermatogenesis by small non-coding RNAs: Role of the germ granule. Semin. Cell Dev. Biol. 29, 84–92.

de Siqueira-Silva, D.H., dos Santos Silva, A.P., Ninhaus-Silveira, A., Veríssimo-Silveira, R., 2015. The effects of temperature and busulfan (Myleran) on the yellowtail tetra *Astyanax altiparanae* (Pisces, Characiformes) spermatogenesis. Theriogenology 84, 1033–1042.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602–607.

Denbow, C.J., Lapins, S., Dietz, N., Scherer, R., Nimchuk, Z.L., Okumoto, S., 2017. Gateway-Compatible CRISPR-Cas9 Vectors and a Rapid Detection by High-Resolution Melting Curve Analysis. Front. Plant Sci. 8, 1171.

Deng, W., Lin, H., 2002. *miwi*, a Murine Homolog of *piwi*, Encodes a Cytoplasmic Protein Essential for Spermatogenesis. Dev. Cell 2, 819–830.

Deniz Koç, N., Yüce, R., 2012. A light- and electron microscopic study of primordial germ cells in the zebra fish (*Danio rerio*). Biol. Res. 45, 331–336.

Deriano, L., Roth, D.B., 2013. Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. Annu. Rev. Genet. 47, 433–455.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., Moineau, S., 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J. Bacteriol. 190, 1390–400.

Dicou, E., Perez-Polo, J.R., 2009. Bax-an emerging role in ectopic cell death. Int. J. Dev. Neurosci. 27, 299–304.

Dixon, D., 1994. Science: On the fish farm, boys are best. (30 July 1994) *New Scientist* Retreived from https://www.newscientist.com/article/mg14319362.700-science-on-the-fish-farm-boys-are-best-/

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H.W., Listgarten, J., Root,

D.E., 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. 34, 184–191.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E., 2014. Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nat. Biotechnol. 32, 1262–1267.

Dohle, G.R., Smit, M., Weber, R.F.A., 2003. Androgens and male fertility. World J. Urol. 21, 341–345.

Doitsidou, M., Reichman-Fried, M., Stebler, J., Köprunner, M., Dörries, J., Meyer, D., Esguerra, C. V., Leung, T., Raz, E., 2002. Guidance of Primordial Germ Cell Migration by the Chemokine SDF-1. Cell 111, 647–659.

Donaldson, E.M., Devlin, R.H., Solar, I.I., Piferrer, F., 1993. The Reproductive Containment of Genetically Altered Salmonids, in: Cloud, O.G., Thorgaard, G.H. (Eds.), Genetic Conservation of Salmonid. Springer US, New York, pp. 113–130.

Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Amacher, S.L., 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat. Biotechnol. 26, 702–8.

Draper, B.W., McCallum, C.M., Moens, C.B., 2007. nanos1 is required to maintain oocyte production in adult zebrafish. Dev. Biol. 305, 589–98.

Duan, J., Feng, G., Chang, P., Zhang, X., Zhou, Q., Zhong, X., Qi, C., Xie, S., Zhao, H., 2015. Germ cell-specific expression of *dead end* (*dnd*) in rare minnow (*Gobiocypris rarus*). Fish Physiol. Biochem. 41, 561–571.

Edvardsen, R.B., Leininger, S., Kleppe, L., Skaftnesmo, K.O., Wargelius, A., 2014. Targeted mutagenesis in Atlantic salmon (*Salmo salar* L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. PLoS One 9, e108622.

Eriksson, D., Harwood, W., Hofvander, P., Jones, H., Rogowsky, P., Stöger, E., Visser, R.G.F., 2018. A Welcome Proposal to Amend the GMO Legislation of the EU. Trends Biotechnol. *In press*. https://doi.org/10.1016/j.tibtech.2018.05.001.

European Parliament, Council of the European Union, 2003. Council Directive of 22 September 2003 on genetically modified food and feed, Regulation (EC) No 1829/2003. Official J. L. 268, 1–23.

Ewen-Campen, B., Schwager, E.E., Extavour, C.G.M., 2010. The molecular machinery of germ line specification. Mol. Reprod. Dev. 77, 3–18.

Extavour, C.G., Akam, M., 2003. Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. Development 130, 5869–5884.

Fan, L., Moon, J., Wong, T.-T., Crodian, J., Collodi, P., 2008. Zebrafish primordial germ cell cultures derived from *vasa*::RFP transgenic embryos. Stem Cells Dev. 17, 585–597.

FAO, 2016. Planning for Aquaculture Diversification: The Importance of Climate Change and Other Drivers: FAO Technical Workshop. FAO fisheries and aquaculture proceedings 47, 51–92.

FAO, 2017. FAO Yearbook. Fishery and Aquaculture Statistics. 2015. FAO, Rome, Italy.

Farlora, R., Hattori-Ihara, S., Takeuchi, Y., Hayashi, M., Octavera, A., Alimuddin, Yoshizaki, G., 2014. Intraperitoneal germ cell transplantation in the Nile tilapia *Oreochromis niloticus*. Mar. Biotechnol. (NY). 16, 309–320.

Fauvel, C., Savoye, O., Dreanno, C., Cosson, J., Suquet, M., 1999. Characteristics of sperm of captive seabass in relation to its fertilization potential. J. Fish Biol. 54, 356–369.

Feng, R., Fang, L., Cheng, Y., He, X., Jiang, W., Dong, R., Shi, H., Jiang, D., Sun, L., Wang, D., 2015. Retinoic acid homeostasis through *aldh1a2* and *cyp26a1* mediates meiotic entry in Nile tilapia (*Oreochromis niloticus*). Sci. Rep. 5, 10131.

Fernández, J.A., Bubner, E.J., Takeuchi, Y., Yoshizaki, G., Wang, T., Cummins, S.F., Elizur, A., 2015. Primordial germ cell migration in the yellowtail kingfish (*Seriola lalandi*) and identification of stromal cell-derived factor 1. Gen. Comp. Endocrinol. 213, 16–23.

Fernandino, J.I., Hattori, R.S., Kishii, A., Strüssmann, C.A., Somoza, G.M., 2012. The Cortisol and Androgen Pathways Cross Talk in High Temperature-Induced Masculinization: The 11β-Hydroxysteroid Dehydrogenase as a Key Enzyme. Endocrinology 153, 6003–6011.

Fletcher, G.L., Shears, M.A., King, M.J., Davies, P.L., Hew, C.L., 1988. Evidence for Antifreeze Protein Gene Transfer in Atlantic Salmon (*Salmo salar*). Can. J. Fish. Aquat. Sci. 45, 352–357.

Forbes, A., Lehmann, R., 1998. *Nanos* and *Pumilio* have critical roles in the development and function of *Drosophila* germline stem cells. Development 125, 679–690.

Fraser, T.W.K., Fjelldal, P.G., Hansen, T., Mayer, I., 2012. Welfare Considerations of Triploid Fish. Rev. Fish. Sci. 20, 192–211.

Fujimoto, T., Nishimura, T., Goto-Kazeto, R., Kawakami, Y., Yamaha, E., Arai, K., 2010. Sexual dimorphism of gonadal structure and gene expression in germ cell-deficient loach, a teleost fish. Proc. Natl. Acad. Sci. 107, 17211–17216.

Fujimura, K., Conte, M.A., Kocher, T.D., 2011. Circular DNA intermediate in the duplication of Nile tilapia *vasa* genes. PLoS One 6, e29477.

Fujimura, K., Kocher, T.D., 2011. Tol2-mediated transgenesis in tilapia (*Oreochromis niloticus*). Aquaculture 319, 342–346.

Fujimura, K., Okada, N., 2007. Development of the embryo, larva and early juvenile of Nile tilapia *Oreochromis niloticus* (Pisces: Cichlidae). Dev. Growth Differ. 49, 301–324.

Gaj, T., Gersbach, C.A., Barbas, C.F., 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31, 397–405.

Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. 109, E2579-E2586.

Gay, S., Bugeon, J., Bouchareb, A., Henry, L., Montfort, J., Cam, A. Le, Bobe, J., Thermes, V., 2018. MicroRNA-202 (miR-202) controls female fecundity by regulating medaka oogenesis. bioRxiv 287359.

Gevers, P., Dulos, J., Boogaart, J.G.M., Timmermans, L.P.M., 1992. A study on cell lineage, especially the germ cell line, in embryos of the teleost fish, *Barbus conchonius*. Roux's Arch. Dev. Biol. 201, 275–283.

Giraldez, A.J., 2010. microRNAs, the cell's Nepenthe: clearing the past during the maternal-to-zygotic transition and cellular reprogramming. Curr. Opin. Genet. Dev. 20, 369–375.

Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., Schier, A.F., 2005. MicroRNAs regulate brain morphogenesis in zebrafish. Science 308, 833–838.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., Schier, A.F., 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312, 75–79.

Göbel, U., Schneider, D.T., Calaminus, G., Haas, R.J., Schmidt, P., Harms, D., 2000. Germ-cell tumors in childhood and adolescence. GPOH MAKEI and the MAHO study groups. Ann. Oncol. 11, 263–271.

Golpour, A., Siddique, M.A.M., Siqueira-Silva, D.H., Pšenička, M., 2016. Induced sterility in fish and its potential and challenges for aquaculture and germ cell transplantation technology: a review. Biologia (Bratisl). 71. 853–864

Goos, H.J.T., Consten, D., 2002. Stress adaptation, cortisol and pubertal development in the male common carp, Cyprinus carpio. Mol. Cell. Endocrinol. 197, 105–116.

Goto, R., Saito, T., Kawakami, Y., Kitauchi, T., Takagi, M., Todo, T., Arai, K., Yamaha, E., 2015. Visualization of primordial germ cells in the fertilized pelagic eggs of the barfin flounder *Verasper moseri*. Int. J. Dev. Biol. 59, 465–470.

Goto, R., Saito, T., Takeda, T., Fujimoto, T., Takagi, M., Arai, K., Yamaha, E., 2012. Germ cells are not the primary factor for sexual fate determination in goldfish. Dev. Biol. 370, 98–109.

Greaves, P., 2012. Chapter 11 - Male Genital Tract, in: Histopathology of Preclinical Toxicity Studies (Fourth Edition). Elsevier, pp. 615–666.

Green, D.R., Reed, J.C., 1998. Mitochondria and apoptosis. Science 281, 1309–12.

Gross-Thebing, T., Yigit, S., Pfeiffer, J., Reichman-Fried, M., Bandemer, J., Ruckert, C., Rathmer, C., Goudarzi, M., Stehling, M., Tarbashevich, K., Seggewiss, J., Raz, E., 2017. The Vertebrate Protein Dead End Maintains Primordial Germ Cell Fate by Inhibiting Somatic Differentiation. Dev. Cell 43, 704–715.e5.

Gruidl, M.E., Smith, P.A., Kuznicki, K.A., McCrone, J.S., Kirchner, J., Roussell, D.L., Strome, S., Bennett, K.L., 1996. Multiple potential germ-line helicases are components of the germ-line-specific P granules of *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. 93, 13837–13842.

Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C., Rebar, E.J., 2010. A Rapid and General Assay for Monitoring Endogenous Gene Modification, in: Methods in Molecular Biology (Clifton, N.J.). pp. 247–256.

Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G.F., Chin, L., 2016. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. Cell Rep. 14, 1555–1566.

Guyon, R., Rakotomanga, M., Azzouzi, N., Coutanceau, J.P., Bonillo, C., D'Cotta, H., Pepey, E., Soler, L., Rodier-Goud, M., D'Hont, A., Conte, M.A., van Bers, N.E., Penman, D.J., Hitte, C., Crooijmans, R.P., Kocher, T.D., Ozouf-Costaz, C.,

Baroiller, J., Galibert, F., 2012. A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. BMC Genomics 13, 222.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., Terns, M.P., 2009. RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. Cell 139, 945–956.

Hamasaki, M., Takeuchi, Y., Yazawa, R., Yoshikawa, S., Kadomura, K., Yamada, T., Miyaki, K., Kikuchi, K., Yoshizaki, G., 2017. Production of Tiger Puffer *Takifugu rubripes* Offspring from Triploid Grass Puffer *Takifugu niphobles* Parents. Mar. Biotechnol. 1–13.

Handeland, S.O., Imsland, A.K., Stefansson, S.O., 2008. The effect of temperature and fish size on growth, feed intake, food conversion efficiency and stomach evacuation rate of Atlantic salmon post-smolts. Aquaculture 283, 36–42.

Hartung, O., Forbes, M.M., Marlow, F.L., 2014. Zebrafish *vasa* is required for germ-cell differentiation and maintenance. Mol. Reprod. Dev. 81, 946–961.

Hashimoto, Y., Maegawa, S., Nagai, T., Yamaha, E., Suzuki, H., Yasuda, K., Inoue, K., 2004. Localized maternal factors are required for zebrafish germ cell formation. Dev. Biol. 268, 152–161.

Hay, B., Jan, L.Y., Jan, Y.N., 1988. A protein component of *Drosophila* polar granules is encoded by *vasa* and has extensive sequence similarity to ATP-dependent helicases. Cell 55, 577–587.

Hayashi, Y., Kobira, H., Yamaguchi, T., Shiraishi, E., Yazawa, T., Hirai, T., Kamei, Y., Kitano, T., 2010. High temperature causes masculinization of genetically female medaka by elevation of cortisol. Mol. Reprod. Dev. 77, 679–686.

Heler, R., Samai, P., Modell, J.W., Weiner, C., Goldberg, G.W., Bikard, D., Marraffini, L.A., 2015. Cas9 specifies functional viral targets during CRISPR–Cas adaptation. Nature 519, 199–202.

Hendel, A., Fine, E.J., Bao, G., Porteus, M.H., 2015. Quantifying on- and off-target genome editing. Trends Biotechnol. 33, 132–140.

Hendel, A., Kildebeck, E.J., Fine, E.J., Clark, J., Punjya, N., Sebastiano, V., Bao, G., Porteus, M.H., 2014. Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. Cell Rep. 7, 293–305.

Herpin, A., Rohr, S., Riedel, D., Kluever, N., Raz, E., Schartl, M., 2007. Specification of primordial germ cells in medaka (*Oryzias latipes*). BMC Dev. Biol. 7, 3.

Heyer, W.-D., Ehmsen, K.T., Liu, J., 2010. Regulation of homologous recombination in eukaryotes. Annu. Rev. Genet. 44, 113–139.

Hickling, C.F., 1960. The Malacca tilapia hybrids. J. Genet. 57, 1–10.

HLPE, 2014. Sustainable fisheries and aquaculture for food security and nutrition. A report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security, Rome, 29 p.

Hofmann, M.-C., 2008. Gdnf signaling pathways within the mammalian spermatogonial stem cell niche. Mol. Cell. Endocrinol. 288, 95–103.

Hogan, B.L., 1996. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes Dev. 10, 1580–1594.

Hooe, M.L., Buck, D.H., Wahl, D.H., 1994. Growth, Survival, and Recruitment of Hybrid Crappies Stocked in Small Impoundments. North Am. J. Fish. Manag. 14, 137–142.

Houston, R.D., 2017. Future directions in breeding for disease resistance in aquaculture species. Rev. Bras. Zootec. 46, 545–551.

Houwing, S., Berezikov, E., Ketting, R.F., 2008. *Zili* is required for germ cell differentiation and meiosis in zebrafish. EMBO J. 27, 2702–2711.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D. V, Blaser, H., Raz, E., Moens, C.B., Plasterk, R.H.A., Hannon, G.J., Draper, B.W., Ketting, R.F., 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129, 69–82.

Hsu, C.-C., Hou, M.-F., Hong, J.-R., Wu, J.-L., Her, G.M., 2010. Inducible Male Infertility by Targeted Cell Ablation in Zebrafish Testis. Mar. Biotechnol. 12, 466–478.

Hsu, P.D., Lander, E.S., Zhang, F., 2014. Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell 157, 1262–1278.

Hu, S.-Y., Lin, P.-Y., Liao, C.-H., Gong, H.-Y., Lin, G.-H., Kawakami, K., Wu, J.-L., 2010. Nitroreductase-mediated gonadal dysgenesis for infertility control of genetically modified zebrafish. Mar. Biotechnol. (NY). 12, 569–578.

Hu, W., Wang, Y., Zhu, Z., 2007. Progress in the evaluation of transgenic fish for possible ecological risk and its containment strategies. Sci. China Ser. C Life Sci. 50, 573–579.

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S., Zhang, B., 2011. Heritable gene targeting in zebrafish using customized TALENs. Nat. Biotechnol. 29, 699–700.

Hulata, G., Wohlfarth, G., Rothbard, S., 1983. Progeny-testing selection of tilapia broodstocks producing all-male hybrid progenies — Preliminary results. Aquaculture 33, 263–268.

Ijiri, S., Kaneko, H., Kobayashi, T., Wang, D.-S., Sakai, F., Paul-Prasanth, B., Nakamura, M., Nagahama, Y., 2008. Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia *Oreochromis niloticus*. Biol. Reprod. 78, 333–341.

Ikenishi, K., Tanaka, T.S., 1997. Involvement of the protein of *Xenopus* vasa homolog (*Xenopus* vasa-like gene 1, XVLG1) in the differentiation of primordial germ cells. Dev. Growth Differ. 39, 625–633.

Irawan, D., 1993. Comparison of egg storage in three tilapia genera (*Tilapia zillii, Sarotherodon galilaeus* and *Oreochromis niloticus*). PG diploma, University of Stirling.

Ito, L.S., Takahashi, C., Yamashita, M., Strüssmann, C.A., 2008. Warm water induces apoptosis, gonadal degeneration, and germ cell loss in subadult pejerrey *Odontesthes bonariensis* (Pisces, Atheriniformes). Physiol. Biochem. Zool. 81, 762–774.

Ito, L.S., Yamashita, M., Strüssmann, C.A., 2003. Histological process and dynamics of germ cell degeneration in pejerrey *Odontesthes bonariensis* larvae and juveniles during exposure to warm water. J. Exp. Zool. Part A Comp. Exp. Biol. 297A, 169–179.

Jain, M., Olsen, H.E., Paten, B., Akeson, M., 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 17, 239.

Jao, L.-E., Wente, S.R., Chen, W., 2013. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl. Acad. Sci. 110, 13904–13909.

Jia, Y., Hikim, A.P.S., Lue, Y.-H., Swerdloff, R.S., Vera, Y., Zhang, X.-S., Hu, Z.-Y., Li, Y.-C., Liu, Y.-X., Wang, C., 2007. Signaling Pathways for Germ Cell Death in Adult Cynomolgus Monkeys (*Macaca fascicularis*) Induced by Mild Testicular Hyperthermia and Exogenous Testosterone Treatment1. Biol. Reprod. 77, 83–92.

Jiang, D.-N., Yang, H.-H., Li, M.-H., Shi, H.-J., Zhang, X.-B., Wang, D.-S., 2016. *gsdf* is a downstream gene of *dmrt1* that functions in the male sex determination pathway of the Nile tilapia. Mol. Reprod. Dev. 83, 497–508.

Jiang, D., Chen, J., Fan, Z., Tan, D., Zhao, J., Shi, H., Liu, Z., Tao, W., Li, M., Wang, D., 2017. CRISPR/Cas9-induced disruption of *wt1a* and *wt1b* reveals their different roles in kidney and gonad development in Nile tilapia. Dev. Biol. 428, 63–73.

Jiang, F., Doudna, J.A., 2017. CRISPR–Cas9 Structures and Mechanisms. Annu. Rev. Biophys. 46, 505–529.

Jiang, Y., 1993. Transgenic fish-gene transfer to increase disease and cold resistance. Aquaculture 111, 31–40.

Jin, Y.H., Davie, A., Migaud, H., 2019. Expression pattern of *nanos, piwil, dnd, vasa* and *pum* genes during ontogenic development in Nile tilapia *Oreochromis niloticus*. Gene 688, 62–70.

Jin, Y.-W., Qu, Y.J., Wang, H., Bai, J.L., Song, F., 2012. Limitation of PCR-RFLP method for the detection of genetic mutations in spinal muscular atrophy. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 29, 34–37.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337, 816–821.

Johnson, A.D., Bachvarova, R.F., Drum, M., Masi, T., 2001. Expression of *Axolotl* DAZL RNA, a Marker of Germ Plasm: Widespread Maternal RNA and Onset of Expression in Germ Cells Approaching the Gonad. Dev. Biol. 234, 402–415.

Johnstone, R., McLay, H.A., Walsingham, M.V., 1991. Production and performance of triploid Atlantic salmon in Scotland. Can. Tech. Rep. Fish. Aquat. Sci. 1789, 15–36.

Kapuscinski, A.R., Patronski, T.J., 2005. Genetic Methods for Biological Control of Non-Native Fish in the Gila River Basin: Final Report to the U.S. Fish and Wildlife Service. University of Minnesota, Institute for Social, Economic and Ecological Sustainability, St. Paul, Minnesota.

KC, R., Srivastava, A., Wilkowski, J.M., Richter, C.E., Shavit, J.A., Burke, D.T., Bielas, S.L., 2016. Detection of nucleotide-specific CRISPR/Cas9 modified alleles using multiplex ligation detection. Sci. Rep. 6, 32048.

Kedde, M., Strasser, M.J., Boldajipour, B., Oude Vrielink, J. a F., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P.M., van Duijse, J., Ørom, U.A., Lund, A.H., Perrakis, A., Raz, E., Agami, R., 2007. RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell 131, 1273–1286.

Khalil, K., Elayat, M., Khalifa, E., Daghash, S., Elaswad, A., Miller, M., Abdelrahman, H., Ye, Z., Odin, R., Drescher, D., Vo, K., Gosh, K., Bugg, W., Robinson, D., Dunham, R., 2017. Generation of Myostatin Gene-Edited Channel Catfish (*Ictalurus punctatus*) via Zygote Injection of CRISPR/Cas9 System. Sci. Rep. 7, 7301.

Khan, M.G.Q., McAndrew, B.J., Penman, D.J., 2015. Validation of clonal line females for sex determination studies in Nile tilapia *Oreochromis niloticus* L. Res. Agric. Livest. Fish. 1(1), 147–158.

Kim, H., Ishidate, T., Ghanta, K.S., Seth, M., Conte, D., Shirayama, M., Mello, C.C., 2014c. A Co-CRISPR Strategy for Efficient Genome Editing in *Caenorhabditis elegans*. Genetics 197, 1069–1080.

Kim, H., Kim, J.-S., 2014. A guide to genome engineering with programmable nucleases. Nat. Rev. Genet. 15, 321–334.

Kim, H., Lee, H., Kim, H., Cho, S., 2009. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. 1279–1288.

Kim, J.M., Kim, D., Kim, S., Kim, J.-S., 2014a. Genotyping with CRISPR-Cas-derived RNA-guided endonucleases. Nat. Commun. 5, 3157.

Kim, S., Kim, D., Cho, S.W., Kim, J.-S., Kim, J.-S., 2014b. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 24, 1012–1019.

Kim, Y.G., Cha, J., Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. 93, 1156–1160.

Kishimoto, K., Washio, Y., Yoshiura, Y., Toyoda, A., Ueno, T., Fukuyama, H., Kato, K., Kinoshita, M., 2018. Production of a breed of red sea bream *Pagrus major* with an increase of skeletal muscle mass and reduced body length by genome editing with CRISPR/Cas9. Aquaculture 495, 415–427.

Klattenhoff, C., Theurkauf, W., 2008. Biogenesis and germline functions of piRNAs. Development 135, 3–9.

Kleppe, L., Wargelius, A., Johnsen, H., Andersson, E., Edvardsen, R.B., 2015. Gonad specific genes in Atlantic salmon (*Salmon salar* L.): Characterization of *tdrd7-2*, *dazl-2*, *piwil1* and *tdrd1* genes. Gene 75, 1–9.

Kloc, M., Jedrzejowska, I., Tworzydlo, W., Bilinski, S.M., 2014. Balbiani body, nuage and sponge bodies--term plasm pathway players. Arthropod Struct. Dev. 43, 341–348.

Knaut, H., Pelegri, F., Bohmann, K., Schwarz, H., Nüsslein-volhard, C., 2000. Zebrafish *vasa* RNA but Not Its Protein Is a Component of the Germ Plasm and Segregates Asymmetrically before Germline Specification. J. Cell Biol. 149, 875–888.

Knaut, H., Steinbeisser, H., Schwarz, H., Nu, C., Planck, M., 2002. An Evolutionary Conserved Region in the *vasa* 3' UTR Targets RNA Translation to the Germ Cells in the Zebrafish. Curr. Biol. 12, 454–466.

Knaut, H., Werz, C., Geisler, R., Nusslein-Volhard, C., 2003. A zebrafish homologue of the chemokine receptor *Cxcr4* is a germ-cell guidance receptor. Nature 279–282.

Kobayashi, T., Kajiura-Kobayashi, H., Nagahama, Y., 2002. Two isoforms of *vasa* homologs in a teleost fish: their differential expression during germ cell differentiation. Mech. Dev. 111, 167–171.

Kobayashi, T., Kajiura-Kobayashi, H., Nagahama, Y., 2000. Differential expression of vasa homologue gene in the germ cells during oogenesis and spermatogenesis in a teleost fish, tilapia, *Oreochromis niloticus*. Mech. Dev. 99, 139–142.

Kobayashi, T., 2010. *In vitro* germ cell differentiation during sex differentiation in a teleost fish. Int. J. Dev. Biol. 54, 105–111.

Kobayashi, Y., Nagahama, Y., Nakamura, M., 2013. Diversity and plasticity of sex determination and differentiation in fishes. Sex Dev. 7, 115–125.

Konno, K., Tashiro, F., 1982. The sterility of rainbow trout (*Salmo gairdneri*) irradiated with cobalt-60 gamma rays. J. Tokyo Univ. Fish. 68, 75–80.

Koonin, E. V, Makarova, K.S., Zhang, F., 2017. Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. 37, 67–78.

Köprunner, M., Thisse, C., Thisse, B., Raz, E., 2001. A zebrafish *nanos*-related gene is essential for the development of primordial germ cells. Genes Dev. 15, 2877–2885.

Kosaka, K., Kawakami, K., Sakamoto, H., Inoue, K., 2007. Spatiotemporal localization of germ plasm RNAs during zebrafish oogenesis. Mech. Dev. 124, 279–289.

Kotaja, N., Bhattacharyya, S.N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., Sassone-Corsi, P., 2006. The chromatoid body of male germ cells: Similarity with processing bodies and presence of Dicer and microRNA pathway components. Proc. Natl. Acad. Sci. 103, 2647–2652.

Kratz, E., Eimon, P.M., Mukhyala, K., Stern, H., Zha, J., Strasser, A., Hart, R., Ashkenazi, A., 2006. Functional characterization of the Bcl-2 gene family in the zebrafish. Cell Death Differ. 13, 1631–1640.

Kroemer, G., Galluzzi, L., Brenner, C., 2007. Mitochondrial membrane permeabilization in cell death. Physiol. Rev. 87, 99–163.

Kumari, J., Flaten, G.E., Škalko-Basnet, N., Tveiten, H., 2017. Molecular transfer to Atlantic salmon ovulated eggs using liposomes. Aquaculture 479, 404–411.

Kuo, M.-W., Wang, S.-H., Chang, J.-C., Chang, C.-H., Huang, L.-J., Lin, H.-H., Yu, A.L.-T., Li, W.-H., Yu, J., 2009. A Novel *puf-A* Gene Predicted from Evolutionary Analysis Is Involved in the Development of Eyes and Primordial Germ-Cells. PLoS One 4, e4980.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., Lin, H., Matsuda, Y., Nakano, T., 2004. *Mili*, a mammalian member of piwi family gene, is essential for spermatogenesis. Development 131, 839–849.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Takamatsu, K., Chuma, S., Kojima-Kita, K., Shiromoto, Y., Asada, N., Toyoda, A., Fujiyama, A., Totoki, Y., Shibata, T., Kimura, T., Nakatsuji, N., Noce, T., Sasaki, H., Nakano, T., 2010. MVH in piRNA processing and gene silencing of retrotransposons. Genes Dev. 24, 887–892.

Kurokawa, H., Saito, D., Nakamura, S., Katoh-Fukui, Y., Ohta, K., Baba, T., Morohashi, K. -i., Tanaka, M., 2007. Germ cells are essential for sexual dimorphism in the medaka gonad. Proc. Natl. Acad. Sci. 104, 16958–16963.

Lacerda, S.M.S.N., Batlouni, S.R., Silva, S.B.G., Homem, C.S.P., França, L.R., 2006. Germ cells transplantation in fish: The Nile-tilapia model, Anim. Reprod. 2, 146-159.

Lacerda, S.M.S.N., Costa, G.M.J., da Silva, M.D.A., Almeida Campos-Junior, P.H., Segatelli, T.M., Peixoto, M.T.D., Resende, R.R., de França, L.R., 2013. Phenotypic characterization and *in vitro* propagation and transplantation of the Nile tilapia (*Oreochromis niloticus*) spermatogonial stem cells. Gen. Comp. Endocrinol. 192, 95–106.

Lahnsteiner, F., Berger, B., Weismann, T., Patzner, R.., 1998. Determination of semen quality of the rainbow trout, *Oncorhynchus mykiss*, by sperm motility, seminal plasma parameters, and spermatozoal metabolism. Aquaculture 163, 163–181.

Lai, F., King, M. Lou, 2013. Repressive translational control in germ cells. Mol. Reprod. Dev. 80, 665–676.

Lambert, Y., Dutil, J.-D., 1997. Can simple condition indices be used to monitor and quantify seasonal changes in the energy reserves of cod (*Gadus morhua*)? Can. J. Fish. Aquat. Sci. 54, 104–112.

Lander, E.S., 2016. The Heroes of CRISPR. Cell 164, 18–28.

Langley, A.R., Smith, J.C., Stemple, D.L., Harvey, S.A., 2014. New insights into the maternal to zygotic transition. Development 141, 3834–3841.

Lauth, X., Buchanan, J.T., 2012. Maternally Induced Sterility in Animals. US20120304323A1.

Layden, M.J., Röttinger, E., Wolenski, F.S., Gilmore, T.D., Martindale, M.Q., 2013. Microinjection of mRNA or morpholinos for reverse genetic analysis in the starlet sea anemone, *Nematostella vectensis*. Nat. Protoc. 8, 924–934.

Ledford, H., 2015. Salmon approval heralds rethink of transgenic animals. Nature 527, 417–418.

Lee, K.H., Yamaguchi, A., Rashid, H., Kadomura, K., Yasumoto, S., Matsuyama, M., 2009. Germ cell degeneration in high-temperature treated pufferfish, *Takifugu rubripes*. Sex Dev. 3, 225–232.

Lehmann, R., 2012. Germline stem cells: origin and destiny. Cell Stem Cell 10, 729–739.

Li, D.-D., Luo, Z., Chen, G.-H., Song, Y.-F., Wei, C.-C., Pan, Y.-X., 2017a. Identification of apoptosis-related genes Bcl2 and Bax from yellow catfish *Pelteobagrus fulvidraco* and their transcriptional responses to waterborne and dietborne zinc exposure. Gene 633, 1–8.

Li, G., Zhang, X., Zhong, C., Mo, J., Quan, R., Yang, J., Liu, D., Li, Z., Yang, H., Wu, Z., 2017b. Small molecules enhance CRISPR/Cas9-mediated homology-directed genome editing in primary cells. Sci. Rep. 7, 8943.

Li, H., Su, B., Qin, G., Ye, Z., Alsaqufi, A., Perera, D., Shang, M., Odin, R., Vo, K., Drescher, D., Robinson, D., Zhang, D., Abass, N., Dunham, R., 2017c. Salt Sensitive Tet-Off-Like Systems to Knockdown Primordial Germ Cell Genes for Repressible Transgenic Sterilization in Channel Catfish, *Ictalurus punctatus*. Mar. Drugs 15, 155.

Li, H., Su, B., Qin, G., Ye, Z., Elaswad, A., Alsaqufi, A., Perera, D.A., Qin, Z., Odin, R., Vo, K., Drescher, D., Robinson, D., Dong, S., Zhang, D., Shang, M., Abass, N., Das, S.K., Bangs, M., Dunham, R.A., 2018. Repressible Transgenic Sterilization in Channel Catfish, *Ictalurus punctatus*, by Knockdown of Primordial Germ Cell Genes with Copper-Sensitive Constructs. Mar. Biotechnol. 20, 324–342.

Li, M., Feng, R., Ma, H., Dong, R., Liu, Z., Jiang, W., Tao, W., Wang, D., 2016. Retinoic acid triggers meiosis initiation via stra8-dependent pathway in Southern catfish, *Silurus meridionalis*. Gen. Comp. Endocrinol. 232, 191–198.

Li, M., Hong, N., Gui, J., Hong, Y., 2012. Medaka *piwi* is Essential for Primordial Germ Cell Migration. Curr. Mol. Med. 12, 1040–1049.

Li, M., Hong, N., Xu, H., Yi, M., Li, C., Gui, J., Hong, Y., 2009. Medaka *vasa* is required for migration but not survival of primordial germ cells. Mech. Dev. 126, 366–381.

Li, M., Shen, Q., Xu, H., Wong, F.M., Cui, J., Li, Z., Hong, N., Wang, L., Zhao, H., Ma, B., Hong, Y., 2011. Differential conservation and divergence of fertility genes *boule* and *dazl* in the rainbow trout. PLoS One 6, e15910.

Li, M., Sun, Y., Zhao, J., Shi, H., Zeng, S., Ye, K., Jiang, D., Zhou, L., Sun, L., Tao, W., Nagahama, Y., Kocher, T.D., Wang, D., 2015. A Tandem Duplicate of Anti-Müllerian Hormone with a Missense SNP on the Y Chromosome Is Essential for Male Sex Determination in Nile Tilapia, *Oreochromis niloticus*. PLOS Genet. 11, e1005678.

Li, M., Wang, D., 2017. Gene editing nuclease and its application in tilapia. Sci. Bull. 62, 165–173.

Li, M., Yang, H., Zhao, J., Fang, L., Shi, H., Li, M., Sun, Y., Zhang, X., Jiang, D., Zhou, L., Wang, D., 2014. Efficient and Heritable Gene Targeting in Tilapia by CRISPR/Cas9. Genetics 197, 591–599.

Li, Q., Shirabe, K., Kuwada, J.Y., 2004. Chemokine signaling regulates sensory cell migration in zebrafish. Dev. Biol. 269, 123–136.

Li, Q., Shirabe, K., Thisse, C., Thisse, B., Okamoto, H., Masai, I., Kuwada, J.Y., 2005. Chemokine signaling guides axons within the retina in zebrafish. J. Neurosci. 25, 1711–1717.

Liang, X., Potter, J., Kumar, S., Ravinder, N., Chesnut, J.D., 2017. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. J. Biotechnol. 241, 136–146.

Lin, F., Liu, Q., Li, M., Li, Z., Hong, N., Li, J., Hong, Y., 2012. Transient and stable GFP expression in germ cells by the vasa regulatory sequences from the red seabream (*Pagrus major*). Int. J. Biol. Sci. 8, 882–890.

Lin, F., Zhao, C.Y., Xu, S.H., Ma, D.Y., Xiao, Z.Z., Xiao, Y.S., Xu, C.A., Liu, Q.H., Li, J., 2013. Germline-specific and sexually dimorphic expression of a *dead end* gene homologue in turbot (*Scophthalmus maximus*). Theriogenology 80, 665–672.

Lin, H., Spradling, A.C., 1997. A novel group of *pumilio* mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. Development 124, 2463–2476.

Lin, S., Staahl, B.T., Alla, R.K., Doudna, J.A., 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife 3, e04766.

Little, D.C., Hulata, G., 2000. Strategies for tilapia seed production, in: Tilapias: Biology and Exploitation. Springer Netherlands, Dordrecht, pp. 267–326.

Liu, J., Carmell, M.A., Rivas, F. V, Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., Hannon, G.J., 2004. Argonaute2 is the catalytic engine of mammalian RNAi. Science 305, 1437–1441.

Liu, L., Chen, P., Wang, M., Li, X., Wang, J., Yin, M., Wang, Y., 2017. C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism. Mol. Cell 65, 310–322.

Liu, L., Hong, N., Xu, H., Li, M., Yan, Y., Purwanti, Y., Yi, M., Li, Z., Wang, L., Hong, Y., 2009. Medaka *dead end* encodes a cytoplasmic protein and identifies embryonic and adult germ cells. Gene Expr. Patterns 9, 541–548.

Lokman, P.M., Harris, B., Kusakabe, M., Kime, D.E., Schulz, R.W., Adachi, S., Young, G., 2002. 11-Oxygenated androgens in female teleosts: prevalence, abundance, and life history implications. Gen. Comp. Endocrinol. 129, 1–12.

Lue, Y.-H., Lasley, B.L., Laughlin, L.S., Swerdloff, R.S., Hikim, A.P.S., Leung, A., Overstreet, J.W., Wang, C., 2002. Mild testicular hyperthermia induces profound transitional spermatogenic suppression through increased germ cell apoptosis in adult cynomolgus monkeys (*Macaca fascicularis*). J. Androl. 23, 799–805.

Lue, Y.-H., Sinha Hikim, A.P., Swerdloff, R.S., Im, P., Taing, K.S., Bui, T., Leung, A., Wang, C., 1999. Single Exposure to Heat Induces Stage-Specific Germ Cell Apoptosis in Rats: Role of Intratesticular Testosterone on Stage Specificity 1. Endocrinology 140, 1709–1717.

Lund, E., Liu, M., Hartley, R.S., Sheets, M.D., Dahlberg, J.E., 2009. Deadenylation of maternal mRNAs mediated by miR-427 in *Xenopus laevis* embryos. RNA 15, 2351–2363.

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y., Xie, Y., Shen, R., Chen, S., Wang, Z., Chen, Y., Guo, J., Chen, L., Zhao, X., Dong, Z., Liu, Y.-G., 2015. A Robust CRISPR/Cas9 System for Convenient, High-

Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. Mol. Plant 8, 1274–1284.

Maclean, N., Rahman, M.A., Sohm, F., Hwang, G., Iyengar, A., Ayad, H., Smith, A., Farahmand, H., 2002. Transgenic tilapia and the tilapia genome. Gene 295, 265–277.

Maegawa, S., Yasuda, K., Inoue, K., 1999. Maternal mRNA localization of zebrafish DAZ-like gene. Mech. Dev. 81, 223–226.

Mahabaleshwar, H., Boldajipour, B., Raz, E., 2008. Killing the messenger: The role of CXCR7 in regulating primordial germ cell migration. Cell Adh. Migr. 2, 69–70.

Majhi, S.K., Hattori, R.S., Rahman, S.M., Suzuki, T., Strüssmann, C.A., 2009. Experimentally induced depletion of germ cells in sub-adult Patagonian pejerrey (*Odontesthes hatcheri*). Theriogenology 71, 1162–1172.

Majhi, S.K., Rasal, A.R., Kushwaha, B., Raizada, S., Kotrbova, A., Rozinek, J., Smith, A., Farahmand, H., 2017. Heat and chemical treatments in adult *Cyprinus carpio* (Pisces cypriniformes) rapidly produce sterile gonads. Anim. Reprod. Sci. 55, 364–371.

Makarova, K.S., Grishin, N. V, Shabalina, S.A., Wolf, Y.I., Koonin, E. V, 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct 1, 7.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J.M., Terns, R.M., Terns, M.P., White, M.F., Yakunin, A.F., Garrett, R.A., van der Oost, J., Backofen, R., Koonin, E. V., 2015. An updated evolutionary classification of CRISPR–Cas systems. Nat. Rev. Microbiol. 13, 722–736.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823–826.

Mallet, J., 2007. Hybrid speciation. Nature 446, 279–283.

Mandich, A., Massari, A., Bottero, S., Marino, G., 2002. Histological and histochemical study of female germ cell development in the dusky grouper *Epinephelus marginatus* (Lowe, 1834). Eur. J. Histochem. 46, 87–100.

Mansfield, E.S., Vainer, M., Enad, S., Barker, D.L., Harris, D., Rappaport, E., Fortina, P., 1996. Sensitivity, reproducibility, and accuracy in short tandem repeat genotyping using capillary array electrophoresis. Genome Res. 6, 893–903.

Marlow, F., 2015. Primordial Germ Cell Specification and Migration. F1000Research 4.

Marraffini, L.A., Sontheimer, E.J., 2008. CRISPR interference limits horizontal gene transfer in *staphylococci* by targeting DNA. Science 322, 1843–1845.

Matta, S.L.P., Vilela, D.A.R., Godinho, H.P., França, L.R., 2002. The Goitrogen 6- n - Propyl-2-Thiouracil (PTU) Given during Testis Development Increases Sertoli and Germ Cell Numbers per Cyst in Fish: The Tilapia (*Oreochromis niloticus*) Model. Endocrinology 143, 970–978.

Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., Ploegh, H.L., 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat. Biotechnol. 33, 538–542.

Maxime, V., 2008. The physiology of triploid fish: current knowledge and comparisons with diploid fish. Fish Fish. 9, 67–78.

McEvoy, T., Stack, M., Keane, B., Barry, T., Sreenan, J., Gannon, F., 1988. The expression of a foreign gene in salmon embryos. Aquaculture 68, 27–37.

Mei, J., Gui, J.-F., 2015. Genetic basis and biotechnological manipulation of sexual dimorphism and sex determination in fish. Sci. China. Life Sci. 58, 124–136.

Meistrich, M.L., Shetty, G. (2003) Suppression of testosterone stimulates recovery of spermatogenesis after cancer treatment. Int J Androl 26(3):141–146.

Mianné, J., Codner, G.F., Caulder, A., Fell, R., Hutchison, M., King, R., Stewart, M.E., Wells, S., Teboul, L., 2017. Analysing the outcome of CRISPR-aided genome editing in embryos: Screening, genotyping and quality control. Methods 121–122, 68–76.

Mickoleit, M., Banisch, T.U., Raz, E., 2011. Regulation of hub mRNA stability and translation by miR430 and the dead end protein promotes preferential expression in zebrafish primordial germ cells. Dev. Dyn. 240, 695–703.

Mishima, Y., Giraldez, A.J., Takeda, Y., Fujiwara, T., Sakamoto, H., Schier, A.F., Inoue, K., 2006. Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. Curr. Biol. 16, 2135–2142.

Miyake, A., Saito, T., Kashiwagi, N., Ando, D., Yamamoto, A., Suzuki, T., Nakatsuji, N., Nakatsuji, T., 2006. Cloning and pattern of expression of the shiro-uo *vasa* gene

during embryogenesis and its roles in PGC development. Int. J. Dev. Biol. 50, 619–625.

Miyasaka, N., Knaut, H., Yoshihara, Y., 2007. Cxcl12/Cxcr4 chemokine signaling is required for placode assembly and sensory axon pathfinding in the zebrafish olfactory system. Development 134, 2459–2468.

Moccia, R.D., Munkittrick, K.R., 1987. Relationship between the fertilization of rainbow trout (*Salmo gairdneri*) eggs and the motility of spermatozoa. Theriogenology 27, 679–688.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., Almendros, C., 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155, 733–740.

Molyneaux, K., Wylie, C., 2004. Primordial germ cell migration. Int. J. Dev. Biol. 48, 537–544.

Muir, W.M., Howard, R.D., 1999. Possible ecological risks of transgenic organism release when transgenes affect mating success: Sexual selection and the Trojan gene hypothesis. Proc. Natl. Acad. Sci. 96, 13853–13856.

Myers, J.M., 1986. Tetraploid induction in *Oreochromis* spp. Aquaculture 57, 281–287.

Myosho, T., Otake, H., Masuyama, H., Matsuda, M., Kuroki, Y., Fujiyama, A., Naruse, K., Hamaguchi, S., Sakaizumi, M., 2012. Tracing the Emergence of a Novel Sex-Determining Gene in Medaka, *Oryzias luzonensis*. Genetics 191, 163–170.

Nagasawa, K., Fernandes, J.M.O., Yoshizaki, G., Miwa, M., Babiak, I., 2013. Identification and migration of primordial germ cells in Atlantic salmon, *Salmo salar*: characterization of *vasa*, *dead end*, and *lymphocyte antigen 75* genes. Mol. Reprod. Dev. 80, 118–131.

Nakamura, M., Kobayashi, T., Chang, X.-T., Nagahama, Y., 1998. Gonadal sex differentiation in teleost fish. J. Exp. Zool. 281, 362–372.

Nakamura, M., Nozu, R., Ijiri, S., Kobayashi, T., Hirai, T., Yamaguchi, Y., Seale, A., Lerner, D.T., Grau, G.E., 2015. Sexual characteristics of high-temperature sterilized male Mozambique tilapia, *Oreochromis mossambicus*. Zool. Lett. 1, 21.

Nóbrega, R.H., Greebe, C.D., van de Kant, H., Bogerd, J., de França, L.R., Schulz, R.W., 2010. Spermatogonial Stem Cell Niche and Spermatogonial Stem Cell Transplantation in Zebrafish. PLoS One 5, e12808.

Norris, D.O., Lopez, K.H., 2011. Hormones and reproduction of vertebrates. Volume 1, Fishes. Academic Press Inc., San Diego.

Nunes, C., Silva, A., Soares, E., Ganias, K., 2011. The Use of Hepatic and Somatic Indices and Histological Information to Characterize the Reproductive Dynamics of Atlantic Sardine *Sardina pilchardus* from the Portuguese Coast. Mar. Coast. Fish. 3, 127–144.

Oatley, J.M., Brinster, R.L., 2012. The germline stem cell niche unit in mammalian testes. Physiol. Rev. 92, 577–595.

Oliver, D., Yuan, S., McSwiggin, H., Yan, W., 2015. Pervasive Genotypic Mosaicism in Founder Mice Derived from Genome Editing through Pronuclear Injection. PLoS One 10, e0129457.

Oltval, Z.N., Milliman, C.L., Korsmeyer, S.J., 1993. Bcl-2 heterodimerizes *in vivo* with a conserved homolog, *Bax*, that accelerates programed cell death. Cell 74, 609–619.

Orthwein, A., Fradet-Turcotte, A., Noordermeer, S.M., Canny, M.D., Brun, C.M., Strecker, J., Escribano-Diaz, C., Durocher, D., 2014. Mitosis Inhibits DNA Double-Strand Break Repair to Guard Against Telomere Fusions. Science 344, 189–193.

Otani, S., Maegawa, S., Inoue, K., Arai, K., Yamaha, E., 2002. The Germ Cell Lineage Identified by *vas* -mRNA during the Embryogenesis in Goldfish. Zoolog. Sci. 19, 519–526.

Ozaki, Y., Higuchi, M., Miura, C., Yamaguchi, S., Tozawa, Y., Miura, T., 2006. Roles of 11β-Hydroxysteroid Dehydrogenase in Fish Spermatogenesis. Endocrinology 147, 5139–5146.

Palaiokostas, C., Bekaert, M., Davie, A., Cowan, M.E., Oral, M., Taggart, J.B., Gharbi, K., McAndrew, B.J., Penman, D.J., Migaud, H., 2013a. Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. BMC Genomics 14, 566.

Palaiokostas, C., Bekaert, M., Khan, M.G.Q., Taggart, J.B., Gharbi, K., McAndrew, B.J., Penman, D.J., 2013b. Mapping and validation of the major sex-determining region in Nile tilapia (*Oreochromis niloticus* L.) Using RAD sequencing. PLoS One 8, e68389.

Palevitch, O., Abraham, E., Borodovsky, N., Levkowitz, G., Zohar, Y., Gothilf, Y., 2010. Cxcl12a-Cxcr4b signaling is important for proper development of the forebrain GnRH system in zebrafish. Gen. Comp. Endocrinol. 165, 262–268.

Pan, C., Ye, L., Qin, L., Liu, X., He, Y., Wang, J., Chen, L., Lu, G., 2016. CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. Sci. Rep. 6, 24765.

Pandit, N.P., Bhandari, R.K., Kobayashi, Y., Nakamura, M., 2015. High temperature-induced sterility in the female Nile tilapia, *Oreochromis niloticus*. Gen. Comp. Endocrinol. 213, 110–117.

Pandit, N.P., Nakamura, M., 2010. Effect of High Temperature on Survival, Growth and Feed Conversion Ratio of Nile Tilapia, *Oreochromis niloticus*. Our Nat. 8, 219–224.

Pankhurst, N.W., Fitzgibbon, Q.P., Pankhurst, P.M., King, H.R., 2008. Habitat-related variation in reproductive endocrine condition in the coral reef damselfish *Acanthochromis polyacanthus*. Gen. Comp. Endocrinol. 155, 386–397.

Parikh, B.A., Beckman, D.L., Patel, S.J., White, J.M., Yokoyama, W.M., 2015. Detailed phenotypic and molecular analyses of genetically modified mice generated by CRISPR-Cas9-mediated editing. PLoS One 10, e0116484.

Peng, J.-X., Xie, J.-L., Zhou, L., Hong, Y.-H., Gui, J.-F., 2009. Evolutionary conservation of *Dazl* genomic organization and its continuous and dynamic distribution throughout germline development in gynogenetic gibel carp. J. Exp. Zool. B. Mol. Dev. Evol. 312, 855–871.

Penman, D.J., Skibinski, D.O.F., Beardmore, J.A., 1987. Survival, growth rate and maturity in triploid tilapia, in: Tiews, K. (Ed.), Proceedings of the World Symposium on Selection, Hybridization, and Genetic Engineering in Acquaculture. Berlin (Germany, F.R.) Heenemann Verlagsgesellschaft, pp. 277–288.

Pfennig, F., Kurth, T., Meissner, S., Standke, A., Hoppe, M., Zieschang, F., Reitmayer, C., Gobel, A., Kretzschmar, G., Gutzeit, H.O., 2012. The social status of the male Nile tilapia (*Oreochromis niloticus*) influences testis structure and gene expression. Reproduction 143, 71–84.

Phelps, R.P., Popma, T.J., 2000. Sex reversal of tilapia, in: Costa-Pierce, B.A., Rakocy, J.E. (Eds.), Tilapia Aquaculture in the Americas, Vol. 2. The World Aquaculture Society, Baton Rouge, Louisiana, United States, pp. 34–59.

Pickering, A.D., Pottinger, T.G., Carragher, J., Sumpter, J.P., 1987. The effects of acute and chronic stress on the levels of reproductive hormones in the plasma of mature male brown trout, *Salmo trutta* L. Gen. Comp. Endocrinol. 68, 249–259.

Piferrer, F., Beaumont, A., Falguière, J.-C., Flajšhans, M., Haffray, P., Colombo, L., 2009. Polyploid fish and shellfish: Production, biology and applications to aquaculture for performance improvement and genetic containment. Aquaculture 293, 125–156.

Pillai, R.S., Chuma, S., 2012. piRNAs and their involvement in male germline development in mice. Dev. Growth Differ. 54, 78–92.

Pinello, L., Canver, M.C., Hoban, M.D., Orkin, S.H., Kohn, D.B., Bauer, D.E., Yuan, G.-C., 2016. Analyzing CRISPR genome-editing experiments with CRISPResso. Nat. Biotechnol. 34, 695–697.

Pinkert, C.A., 1994. Transgenic animal technology: a laboratory handbook. Academic Press Inc., San Diego

Pongthana, N., Penman, D.J., Baoprasertkul, P., Hussain, M.G., Shahidul Islam, M., Powell, S.F., McAndrew, B.J., 1999. Monosex female production in the silver barb (*Puntius gonionotus* Bleeker). Aquaculture 173, 247–256.

Presslauer, C., Nagasawa, K., Dahle, D., Babiak, J., Fernandes, J.M.O., Babiak, I., 2014. Induced Autoimmunity against Gonadal Proteins Affects Gonadal Development in Juvenile Zebrafish. PLoS One 9, e114209.

Presslauer, C., Nagasawa, K., Fernandes, J.M.O., Babiak, I., 2012. Expression of *vasa* and *nanos3* during primordial germ cell formation and migration in Atlantic cod (*Gadus morhua* L.). Theriogenology 78, 1262–1277.

Price, E.P., Smith, H., Huygens, F., Giffard, P.M., 2007. High-resolution DNA melt curve analysis of the clustered, regularly interspaced short-palindromic-repeat locus of Campylobacter jejuni. Appl. Environ. Microbiol. 73, 3431–3436.

Prykhozhij, S. V, Steele, S.L., Razaghi, B., Berman, J.N., 2017. A rapid and effective method for screening, sequencing and reporter verification of engineered frameshift mutations in zebrafish. Dis. Model. Mech. 10, 811–822.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183.

Qin, Z., Li, Y., Su, B., Cheng, Q., Ye, Z., Perera, D.A., Fobes, M., Shang, M., Dunham, R.A., 2016. Editing of the Luteinizing Hormone Gene to Sterilize Channel Catfish, *Ictalurus punctatus*, Using a Modified Zinc Finger Nuclease Technology with Electroporation. Mar. Biotechnol. 18, 255–263.

Qiu, Y., Sun, S., Charkraborty, T., Wu, L., Sun, L., Wei, J., Nagahama, Y., Wang, D., Zhou, L. (2015) Figla Favors Ovarian Differentiation by Antagonizing Spermatogenesis in a Teleosts, Nile Tilapia (*Oreochromis niloticus*). PLoS ONE 10(4): e0123900.

Rahman, M.A., Maclean, N., 1992. Production of Transgenic Tilapia (*Oreochromis niloticus*) by One-Cell-Stage Microinjection. Aquaculture 105, 219–232.

Rajpert-De Meyts, E., 2006. Developmental model for the pathogenesis of testicular carcinoma in situ: genetic and environmental aspects. Hum. Reprod. Update 12, 303–323.

Rakocy, J.E., 2005. Cultured Aquatic Species Information Programme. *Oreochromis niloticus*. [WWW Document]. FAO Fish. Aquac. Dep. [online]. URL http://www.fao.org/fishery/culturedspecies/Oreochromis_niloticus/en#tcNA00EA (accessed 3.16.18).

Ramlee, M.K., Yan, T., Cheung, A.M.S., Chuah, C.T.H., Li, S., 2015. High-throughput genotyping of CRISPR/Cas9-mediated mutants using fluorescent PCR-capillary gel electrophoresis. Sci. Rep. 5, 15587.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308.

Ratajczak, M.Z., Zuba-Surma, E., Kucia, M., Reca, R., Wojakowski, W., Ratajczak, J., 2006. The pleiotropic effects of the SDF-1-CXCR4 axis in organogenesis, regeneration and tumorigenesis. Leukemia 20, 1915–1924.

Raz, E., 2003. Primordial germ-cell development: the zebrafish perspective. Nat. Rev. Genet. 4, 690–700.

Reed, G.H., Kent, J.O., Wittwer, C.T., 2007. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. Pharmacogenomics 8, 597–608.

Renaud, J., Boix, C., Charpentier, M., De Cian, A., Cochennec, J., Duvernois-Berthet, E., Perrouault, L., Tesson, L., Edouard, J., Thinard, R., Cherifi, Y., Menoret, S., Fontanière, S., de Crozé, N., Fraichard, A., Sohm, F., Anegon, I., Concordet, J., Giovannangeli, C., 2016. Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. Cell Rep. 14, 2263–2272.

Reva, B., Antipin, Y., Sander, C., 2011. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 39, e118–e118.

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: The European Molecular Biology Open Software Suite. Trends in Genetics 16(6), 276—277.

Richards, A., Thompson, J.T., 1921. The Migration of the Primary Sex-Cells of *Fundulus Heteroclitus*. Biol. Bull. 40, 325–348.

Richardson, B.E., Lehmann, R., 2010. Mechanisms guiding primordial germ cell migration: strategies from different organisms. Nat. Rev. Mol. Cell Biol. 11, 37–49.

Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., Corn, J.E., 2016. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat. Biotechnol. 34, 339–344.

Rodgers, K., McVey, M., 2016. Error-Prone Repair of DNA Double-Strand Breaks. J. Cell. Physiol. 231, 15–24.

Rosen, J.N., Sweeney, M.F., Mably, J.D., 2009. Microinjection of zebrafish embryos to analyze gene function. J. Vis. Exp. 1–5.

Rucker, E.B., Dierisseau, P., Wagner, K.U., Garrett, L., Wynshaw-Boris, A., Flaws, J.A., Hennighausen, L., 2000. Bcl-x and Bax regulate mouse primordial germ cell survival and apoptosis during embryogenesis. Mol. Endocrinol. 14, 1038–1052.

Ruiz-Verdugo, C.A., Ramírez, J.L., Allen, S.K., Ibarra, A.M., 2000. Triploid catarina scallop (*Argopecten ventricosus* Sowerby II, 1842): growth, gametogenesis, and suppression of functional hermaphroditism. Aquaculture 186, 13–32.

Rurangwa, E., Volckaert, F.A.M., Huyskens, G., Kime, D.E., Ollevier, F., 2001. Quality control of refrigerated and cryopreserved semen using computer-assisted sperm analysis (CASA), viable staining and standardized fertilization in African catfish (*Clarias gariepinus*). Theriogenology 55, 751–769.

Sada, A., Suzuki, A., Suzuki, H., Saga, Y., 2009. The RNA-binding protein NANOS2 is required to maintain murine spermatogonial stem cells. Science 325, 1394–1398.

Saito, T., Fujimoto, T., Maegawa, S., Inoue, K., Tanaka, M., Arai, K., Yamaha, E., 2006. Visualization of primordial germ cells *in vivo* using GFP-*nos1* 3'UTR mRNA. Int. J. Dev. Biol. 50, 691–699.

Saito, T., Goto-Kazeto, R., Kawakami, Y., Nomura, K., Tanaka, H., Adachi, S., Arai, K., Yamaha, E., 2011. The mechanism for primordial germ-cell migration is conserved between Japanese eel and zebrafish. PLoS One 6.

Saito, T., Otani, S., Fujimoto, T., Suzuki, T., Nakatsuji, T., Arai, K., Yamaha, E., 2004. The germ line lineage in ukigori, *Gymnogobius* species (Teleostei: Gobiidae) during embryonic development. Int. J. Dev. Biol. 48, 1079–1085.

Saitou, M., Yamaji, M., 2012. Primordial germ cells in mice. Cold Spring Harb. Perspect. Biol. 4, a008375.

Salsman, J., Dellaire, G., 2017. Precision genome editing in the CRISPR era. Biochem. Cell Biol. 95, 187–201.

Samarut, É., Lissouba, A., Drapeau, P., 2016. A simplified method for identifying early CRISPR-induced indels in zebrafish embryos using High Resolution Melting analysis. BMC Genomics 17, 547.

Santos, A.C., Lehmann, R., 2004. Germ Cell Specification and Migration in *Drosophila* and beyond. Curr. Biol. 14, R578–R589.

Satoh, N., Egami, N., 1972. Sex differentiation of germ cells in the teleost, Oryzias latipes, during noraml embryonic development. J. Embryol. Exp. Morphol. 28, 385–395.

Sawamura, R., Osafune, N., Murakami, T., Furukawa, F., Kitano, T., 2017. Generation of biallelic F0 mutants in medaka using the CRISPR/Cas9 system. Genes to Cells 22, 756–763.

Saxe, J.P., Lin, H., 2011. Small noncoding RNAs in the germline. Cold Spring Harb. Perspect. Biol. 3, a002717.

Schneider, D.T., Schuster, A.E., Fritsch, M.K., Hu, J., Olson, T., Lauer, S., Göbel, U., Perlman, E.J., 2001. Multipoint imprinting analysis indicates a common precursor cell for gonadal and nongonadal pediatric germ cell tumors. Cancer Res. 61, 7268–7276.

Schreck, C.B., 2009. Stress and fish reproduction: The roles of allostasis and hormesis. Gen. Comp. Endocrinol. 165, 549–556.

Schulz, R.W., de França, L.R., Lareyre, J.-J., Le Gac, F., LeGac, F., Chiarini-Garcia, H., Nobrega, R.H., Miura, T., 2010. Spermatogenesis in fish. Gen. Comp. Endocrinol. 165, 390–411.

Scott, A.G., Penman, D.J., Beardmore, J.A., Skibinski, D.O.F., 1989. The 'YY' supermale in *Oreochromis niloticus* (L.) and its potential in aquaculture. Aquaculture 78, 237–251.

Searle, P.F., Chen, M.-J., Hu, L., Race, P.R., Lovering, A.L., Grove, J.I., Guise, C., Jaberipour, M., James, N.D., Mautner, V., Young, L.S., Kerr, D.J., Mountain, A.,

White, S.A., Hyde, E.I., 2004. Nitroreductase: a prodrug-activating enzyme for cancer gene therapy. Clin. Exp. Pharmacol. Physiol. 31, 811–816.

Seeger, C., Sohn, J.A., 2016. Complete Spectrum of CRISPR/Cas9-induced Mutations on HBV cccDNA. Mol. Ther. 24, 1258–1266.

Sekido, R., Lovell-Badge, R., 2008. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature 453, 930–934.

Shao, S., Ren, C., Liu, Z., Bai, Y., Chen, Z., Wei, Z., Wang, X., Zhang, Z., Xu, K., 2017. Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing *Saccharomyces cerevisiae* Rad52. Int. J. Biochem. Cell Biol. 92, 43–52.

Shao, Y., Guan, Y., Wang, L., Qiu, Z., Liu, M., Chen, Y., Wu, L., Li, Y., Ma, X., Liu, M., Li, D., 2014. CRISPR/Cas-mediated genome editing in the rat via direct injection of one-cell embryos. Nat. Protoc. 9, 2493–2512.

Shigeta, M., Sakane, Y., Iida, M., Suzuki, M., Kashiwagi, K., Kashiwagi, A., Fujii, S., Yamamoto, T., Suzuki, K.T., 2016. Rapid and efficient analysis of gene function using CRISPR-Cas9 in *Xenopus tropicalis* founders. Genes to Cells 21, 755–771.

Shimizu, Y., Shibata, N., Yamashita, M., 1997. Spermiogenesis without preceding meiosis in the hybrid medaka between *Oryzias latipes* and *O. curvinotus*. J. Exp. Zool. 279, 102–112.

Shinomiya, A., Tanaka, M., Kobayashi, T., Nagahama, Y., Hamaguchi, S., 2000. The *vasa*-like gene, *olvas*, identifies the migration path of primordial germ cells during embryonic body formation stage in the medaka, *Oryzias latipes*. Dev. Growth Differ. 42, 317–326.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., Koonin, E.V., 2015. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol. Cell 60, 385–397.

Shui, B., Hernandez Matias, L., Guo, Y., Peng, Y., 2016. The Rise of CRISPR/Cas for Genome Editing in Stem Cells. Stem Cells Int. 2016, 8140168.

Siegfried, K.R., Nüsslein-Volhard, C., 2008. Germ line control of female sex determination in zebrafish. Dev. Biol. 324, 277–287.

Simon, B., Kirkpatrick, J.P., Eckhardt, S., Reuter, M., Rocha, E.A., Andrade-Navarro, M.A., Sehr, P., Pillai, R.S., Carlomagno, T., 2011. Recognition of 2'-O-methylated 3'-end of piRNA by the PAZ domain of a Piwi protein. Structure 19, 172–180.

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., Siksnys, V., 2011. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J. 30, 1335–1342.

Skakkebaek, N.E., 1972. Possible carcinoma-in-situ of the testis. Lancet 2, 516–517.

Škugor, A., Slanchev, K., Torgersen, J.S., Tveiten, H., Andersen, Ø., 2014a. Conserved Mechanisms for Germ Cell-Specific Localization of *nanos3* Transcripts in Teleost Species with Aquaculture Significance. Mar. Biotechnol. 16, 256–264.

Škugor, A., Tveiten, H., Krasnov, A., Andersen, Ø., 2014b. Knockdown of the germ cell factor Dead end induces multiple transcriptional changes in Atlantic cod (*Gadus morhua*) hatchlings. Anim. Reprod. Sci. 144, 129–137.

Slanchev, K., Stebler, J., Goudarzi, M., Cojocaru, V., Weidinger, G., Raz, E., 2009. Control of Dead end localization and activity--implications for the function of the protein in antagonizing miRNA function. Mech. Dev. 126, 270–277.

Smargon, A.A., Cox, D.B.T., Pyzocha, N.K., Zheng, K., Slaymaker, I.M., Gootenberg, J.S., Abudayyeh, O.A., Essletzbichler, P., Shmakov, S., Makarova, K.S., Koonin, E. V, Zhang, F., 2017. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. Mol. Cell 65, 618–630.e7.

Smith, L.D., 1966. The role of a "germinal plasm" in the formation of primordial germ cells in Rana pipiens. Dev. Biol. 14, 330–347.

Socher, S.A., Yin, Y., Dewolf, W.C., Morgentaler, A., 1997. Temperature-Mediated Germ Cell Loss in the Testis is Associated With Altered Expression of the Cell-Cycle Regulator *p53*. J. Urol. 157, 1986–1989.

Soler, L., Conte, M.A., Katagiri, T., Howe, A.E., Lee, B.-Y., Amemiya, C., Stuart, A., Dossat, C., Poulain, J., Johnson, J., Di Palma, F., Lindblad-Toh, K., Baroiller, J.-F., D'Cotta, H., Ozouf-Costaz, C., Kocher, T.D., 2010. Comparative physical maps derived from BAC end sequences of tilapia (*Oreochromis niloticus*). BMC Genomics 11, 636.

Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y.E., Zhang, J., 2016. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. Nat. Commun. 7, 10548.

Soria, F.N., Strüssmann, C.A., Miranda, L.A., 2008. High water temperatures impair the reproductive ability of the pejerrey fish *Odontesthes bonariensis*: effects on the hypophyseal-gonadal axis. Physiol. Biochem. Zool. 81, 898–905.

Stoss, J., Holtz, W., 1983. Successful storage of chilled rainbow trout (*Salmo gairdneri*) spermatozoa for up to 34 days. Aquaculture 31, 269–274.

Strasser, M.J., Mackenzie, N.C., Dumstrei, K., Nakkrasae, L.-I., Stebler, J., Raz, E., 2008. Control over the morphology and segregation of Zebrafish germ cell granules during embryonic development. BMC Dev. Biol. 8, 58.

Strome, S., Updike, D., 2015. Specifying and protecting germ cell fate. Nat. Rev. Mol. Cell Biol. 16, 406–416.

Strüssmann, C.A., Saito, T., Takashima, F., 1998. Heat-induced Germ Cell Deficiency in the Teleosts *Odontesthes bonariensis* and *Patagonina hatcheri*. Comp. Biochem. Physiol. Part A Mol. Integr. Physiol. 119, 637–644.

Su, B., Shang, M., Grewe, P.M., Patil, J.G., Peatman, E., Perera, D.A., Cheng, Q., Li, C., Weng, C.-C., Li, P., Liu, Z., Dunham, R.A., 2015. Suppression and restoration of primordial germ cell marker gene expression in channel catfish, *Ictalurus punctatus*, using knockdown constructs regulated by copper transport protein gene promoters: Potential for reversible transgenic sterilization. Theriogenology 84, 1499–512.

Sun, Z.-H., Wang, Y., Lu, W.-J., Li, Z., Liu, X.-C., Li, S.-S., Zhou, L., Gui, J.-F., 2017. Divergent Expression Patterns and Function Implications of Four nanos Genes in a Hermaphroditic Fish, *Epinephelus coioides*. Int. J. Mol. Sci. 18, 685.

Suzuki, H., Sada, A., Yoshida, S., Saga, Y., 2009. The heterogeneity of spermatogonia is revealed by their topology and expression of marker proteins including the germ cell-specific proteins Nanos2 and Nanos3. Dev. Biol. 336, 222–231.

Table 1. Conversion table for calculating microinjection volumes, 2010. Cold Spring Harb. Protoc. 2010, pdb.tab195537.

Tada, H., Mochii, M., Orii, H., Watanabe, K., 2012. Ectopic formation of primordial germ cells by transplantation of the germ plasm: direct evidence for germ cell determinant in *Xenopus*. Dev. Biol. 371, 86–93.

Tanaka, S.S., Toyooka, Y., Akasu, R., Katoh-Fukui, Y., Nakahara, Y., Suzuki, R., Yokoyama, M., Noce, T., 2000. The mouse homolog of *Drosophila* Vasa is required for the development of male germ cells. Genes Dev. 14, 841–853.

Tang, W.W.C., Kobayashi, T., Irie, N., Dietmann, S., Surani, M.A., 2016. Specification and epigenetic programming of the human germ line. Nat. Rev. Genet. 17, 585–600.

Taylor, J.F., Sambraus, F., Mota-Velasco, J., Guy, D.R., Hamilton, A., Hunter, D., Corrigan, D., Migaud, H., 2013. Ploidy and family effects on Atlantic salmon (*Salmo salar*) growth, deformity and harvest quality during a full commercial production cycle. Aquaculture 410–411, 41–50.

Tesarik, J., Greco, E., Cohen-Bacrie, P., Mendoza, C., 1998. Germ cell apoptosis in men with complete and incomplete spermiogenesis failure. Mol. Hum. Reprod. 4, 757–762.

Theusch, E. V, Brown, K.J., Pelegri, F., 2006. Separate pathways of RNA recruitment lead to the compartmentalization of the zebrafish germ plasm. Dev. Biol. 292, 129–141.

Thomas, H.R., Percival, S.M., Yoder, B.K., Parant, J.M., 2014. High-Throughput Genome Editing and Phenotyping Facilitated by High Resolution Melting Curve Analysis. PLoS One 9, e114632.

Thorpe, J.E., Talbot, C., Miles, M.S., 1987. Irradiation of Atlantic salmon eggs to overcome early maturity when selecting for high growth rate, in: Tiews, K. (Ed.), Proceedings of the World Symposium on Selection, Hybridization, and Genetic Engineering in Acquaculture. Berlin (Germany, F.R.) Heenemann Verlagsgesellschaft, Berlin, pp. 361–374.

Tolia, N.H., Joshua-Tor, L., 2007. Slicer and the Argonautes. Nat. Chem. Biol. 3, 36–43.

Trubiroha, A., Gillotay, P., Giusti, N., Gacquer, D., Libert, F., Lefort, A., Haerlingen, B., De Deken, X., Opitz, R., Costagliola, S., 2018. A Rapid CRISPR/Cas-based Mutagenesis Assay in Zebrafish for Identification of Genes Involved in Thyroid Morphogenesis and Function. Sci. Rep. 8, 5647.

Tse, M.Y., Ashbury, J.E., Zwingerman, N., King, W.D., Taylor, S.A., Pang, S.C., 2011. A refined, rapid and reproducible high resolution melt (HRM)-based method suitable for quantification of global LINE-1 repetitive element methylation. BMC Res. Notes 4, 565.

Tsuda, M., Sasaoka, Y., Kiso, M., Abe, K., Haraguchi, S., Kobayashi, S., Saga, Y., 2003. Conserved role of Nanos proteins in germ cell development. Science 301, 1239–1241.

Tsunekawa, N., Naito, M., Sakai, Y., Nishida, T., Noce, T., 2000. Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells. Development 127, 2741–2750.

Tvedt, H.B., Benfey, T.J., Martin-Robichaud, D.J., Power, J., 2001. The relationship between sperm density, spermatocrit, sperm motility and fertilization success in Atlantic halibut, *Hippoglossus hippoglossus*. Aquaculture 194, 191–200.

Uchida, D., Yamashita, M., Kitano, T., Iguchi, T., 2004. An aromatase inhibitor or high water temperature induce oocyte apoptosis and depletion of P450 aromatase activity in the gonads of genetic female zebrafish during sex-reversal. Comp. Biochem. Physiol. A. Mol. Integr. Physiol. 137, 11–20.

Ueno, T., Tanaka, Y.O., Nagata, M., Tsunoda, H., Anno, I., Ishikawa, S., Kawai, K., Itai, Y., 2004. Spectrum of germ cell tumors: from head to toe. Radiographics 24, 387–404.

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Volume I: Comprehensive Tables (ST/ESA/SER.A/399)

Vainio, S., Heikkilä, M., Kispert, A., Chin, N., McMahon, A.P., 1999. Female development in mammals is regulated by Wnt-4 signalling. Nature 397, 405–409.

Valentijn, A.J., Upton, J.-P., Bates, N., Gilmore, A.P., 2008. Bax targeting to mitochondria occurs via both tail anchor-dependent and -independent mechanisms. Cell Death Differ. 15, 1243–1254.

van Winkoop, A., Booms, G.H.R., Dulos, G.J., Timmermans, L.P.M., 1992. Ultrastructural changes in primordial germ cells during early gonadal development of the common carp (*Cyprinus carpio* L., teleostei). Cell Tissue Res. 267, 337–346.

Vladic, T., Jatrvi, T., 1997. Sperm motility and fertilization time span in Atlantic salmon and brown trout-the effect of water temperature. J. Fish Biol. 50, 1088–1093.

Vouillot, L., Thélie, A., Pollet, N., 2015. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. G3 (Bethesda). 5, 407–415.

Wajant, H., 2002. The Fas Signaling Pathway: More Than a Paradigm. Science 296, 1635–1636.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 153, 910–918.

Wang, K., Mei, D.Y., Liu, Q.N., Qiao, X.H., Ruan, W.M., Huang, T., Cao, G.S., 2015. Research of methods to detect genomic mutations induced by CRISPR/Cas systems. J. Biotechnol. 214, 128–132.

Wang, K., Zhang, H., Hu, Q., Shao, C., Chen, S., 2014a. Expression and purification of half-smooth tongue sole (*Cynoglossus semilaevis*) CSDAZL protein. Protein Expr. Purif. 102, 8–12.

Wang, Q., Lu, Y., Xin, Y., Wei, L., Huang, S., Xu, J., 2016. Genome editing of model oleaginous microalgae *Nannochloropsis* spp. by CRISPR/Cas9. Plant J. 88, 1071–1081.

Wang, T., Wei, J.J., Sabatini, D.M., Lander, E.S., 2014b. Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–84.

Wargelius, A., Leininger, S., Skaftnesmo, K.O., Kleppe, L., Andersson, E., Taranger, G.L., Schulz, R.W., Edvardsen, R.B., 2016. Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. Sci. Rep. 6, 21284.

Waring, M.J., 1965. Complex formation between ethidium bromide and nucleic acids. J. Mol. Biol. 13, 269–282.

Weidinger, G., Stebler, J., Slanchev, K., Dumstrei, K., Wise, C., Lovell-Badge, R., Thisse, C., Thisse, B., Raz, E., 2003. *dead end*, a Novel Vertebrate Germ Plasm Component, Is Required for Zebrafish Primordial Germ Cell Migration and Survival. Curr. Biol. 13, 1429–1434.

Weidinger, G., Wolke, U., Köprunner, M., Klinger, M., Raz, E., 1999. Identification of tissues and patterning events required for distinct steps in early migration of zebrafish primordial germ cells. Development 126, 5295–5307.

Whittle, C.A., Extavour, C.G., 2017. Causes and evolutionary consequences of primordial germ-cell specification mode in. Proc.Natl.Acad.Sci.U.S.A.

Wilson, B.E., Netzloff, M.L., 1983. Primary testicular abnormalities causing precocious puberty Leydig cell tumor, Leydig cell hyperplasia, and adrenal rest tumor. Ann. Clin. Lab. Sci. 13, 315–320.

Wolfe, S.A., Nekludova, L., Pabo, C.O., 2000. DNA Recognition by Cys 2 His 2 Zinc Finger Proteins. Annu. Rev. Biophys. Biomol. Struct. 29, 183–212.

Wolke, U., Weidinger, G., Ko, M., Chemistry, M.B., Köprunner, M., Raz, E., 2002. Multiple Levels of Posttranscriptional Control Lead to Germ Line-Specific Gene Expression in the Zebrafish. Curr. Biol. 12, 289–294.

Wong, T.T., Collodi, P., 2013. Inducible Sterilization of Zebrafish by Disruption of Primordial Germ Cell Migration. PLoS One 8, e68455.

Wong, T.T., Zohar, Y., 2015a. Production of reproductively sterile fish: A mini-review of germ cell elimination technologies. Gen. Comp. Endocrinol. 221, 3–8.

Wong, T.T., Zohar, Y., 2015b. Production of reproductively sterile fish by a non-transgenic gene silencing technology. Sci. Rep. 5, 15822.

Woodhead, A.D., Setlow, R.B., 1980. Response to and recovery from acute sublethal gamma-radiation in the Amazon molly, *Poecilia formosa*, in: Egami, M. (Ed.), Radiation Effects on Aquatic Organisms. Japan Scientific Societies Press/University Park Press, Tokyo, pp. 171–182.

Wootton, R.J., Evans, G.W., Mills, L., 1978. Annual cycle in female Three-spined sticklebacks (*Gasterosteus aculeatus* L.) from an upland and lowland population. J. Fish Biol. 12, 331–343.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S., Jaenisch, R., Zhang, F., Sharp, P.A., 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat. Biotechnol. 32, 670–676.

Xiao, J., Zhou, Y., Luo, Y., Zhong, H., Huang, Y., Zhang, Y., Luo, Z., Ling, Z., Zhang, M., Gan, X., 2013. Suppression effect of LHRH-A and hCG on *Piwi* expression in testis of Nile tilapia *Oreochromis niloticus*. Gen. Comp. Endocrinol. 189, 43–50.

Xie, Q.-P., He, X., Sui, Y.-N., Chen, L.-L., Sun, L.-N., Wang, D.-S., 2016. Haploinsufficiency of SF-1 Causes Female to Male Sex Reversal in Nile Tilapia, *Oreochromis niloticus*. Endocrinology en.2015-2049.

Xie, T., 2008. Germline stem cell niches, StemBook. Harvard Stem Cell Institute, Cambridge.

Xu, H., Gui, J., Hong, Y., 2005. Differential expression ofvasa RNA and protein during spermatogenesis and oogenesis in the gibel carp (*Carassius auratus gibelio*), a bisexually and gynogenetically reproducing vertebrate. Dev. Dyn. 233, 872–882.

Xu, H., Li, M., Gui, J., Hong, Y., 2010a. Fish germ cells. Sci. China Life Sci. 53, 435–446.

Xu, H., Li, M., Gui, J., Hong, Y., 2007. Cloning and expression of medaka *dazl* during embryogenesis and gametogenesis. Gene Expr. Patterns 7, 332–338.

Xu, Y., Wang, H., Zhou, J., Lei, Y., Zhou, Y., Yang, Q., Ye, D., Li, W., Deng, F., 2010b. Zebrafish Nanos interacts with and regulates the phosphorylation of Mylz2. Biochimie 92, 1812–1817.

Yamaha, E., Goto-Kazeto, R., Saito, T., Kawakami, Y., Fujimoto, T., Adachi, S., Arai, K., 2010. Primordial germ cell in teleost fish with special references to its specification and migration. J. Appl. Ichthyol. 26, 816–822.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., Jaenisch, R., 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 154, 1370–1379.

Yang, Z., Steentoft, C., Hauge, C., Hansen, L., Thomsen, A.L., Niola, F., Vester-Christensen, M.B., Frödin, M., Clausen, H., Wandall, H.H., Bennett, E.P., 2015. Fast and sensitive detection of indels induced by precise gene targeting. Nucleic Acids Res. 43, e59–e59.

Yano, A., Suzuki, K., Yoshizaki, G., 2008. Flow-cytometric isolation of testicular germ cells from rainbow trout (*Oncorhynchus mykiss*) carrying the green fluorescent protein gene driven by trout vasa regulatory regions. Biol. Reprod. 78, 151–158.

Ye, H., Chen, X., Wei, Q., Zhou, L., Liu, T., Gui, J., Li, C., Cao, H., 2012. Molecular and expression characterization of a nanos1 homologue in Chinese sturgeon, *Acipenser sinensis*. Gene 511, 285–292.

Yen, P.H., 2004. Putative biological functions of the DAZ family. Int. J. Androl. 27, 125–129.

Yen, S.-T., Zhang, M., Deng, J.M., Usman, S.J., Smith, C.N., Parker-Thornburg, J., Swinton, P.G., Martin, J.F., Behringer, R.R., 2014. Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. Dev. Biol. 393, 3–9.

Yin, Y., Hawkins, K.L., Dewolf, W.C., Morgentaler, A., 1997. Heat Stress Causes Testicular Germ Cell Apoptosis in Adult Mice. J. Androl. 18, 159–165.

Yoon, C., Kawakami, K., Hopkins, N., 1997. Zebrafish *vasa* homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. Development 124, 3157–3165.

Yoshida, S., Sukeno, M., Nabeshima, Y.-I., 2007. A vasculature-associated niche for undifferentiated spermatogonia in the mouse testis. Science 317, 1722–1726.

Yoshikawa, H., Morishima, K., Fujimoto, T., Arias-Rodriguez, L., Yamaha, E., Arai, K., 2008. Ploidy manipulation using diploid sperm in the loach, *Misgurnus anguillicaudatus*: a review. J. Appl. Ichthyol. 24, 410–414.

Yoshikawa, H., Xu, D., Ino, Y., Yoshino, T., Hayashida, T., Wang, J., Yazawa, R., Yoshizaki, G., Takeuchi, Y., 2018. Hybrid Sterility in Fish Caused by Mitotic Arrest of Primordial Germ Cells. Genetics 209, 507–521.

Yoshizaki, G., Tago, Y., Takeuchi, Y., Sawatari, E., Kobayashi, T., Takeuchi, T., 2005. Green fluorescent protein labeling of primordial germ cells using a nontransgenic method and its application for germ cell transplantation in salmonidae. Biol. Reprod. 73, 88–93.

Yoshizaki, G., Takeuchi, Y., Sakatani, S., Takeuchi, T., 2000. Germ cell-specific expression of green fluorescent protein in transgenic rainbow trout under control of the rainbow trout vasa-like gene promoter. Int. J. Dev. Biol. 44, 323–6.

Youngren, K.K., Coveney, D., Peng, X., Bhattacharya, C., Schmidt, L.S., Nickerson, M.L., Lamb, B.T., Deng, J.M., Richard, R., Capel, B., Rubin, E.M., Nadeau, J.H., Matin, A., 2005. The Ter mutation in the *dead end* gene causes germ cell loss and testicular germ cell tumours. Nature 435, 360–364.

Yu, C., Zhang, Y., Yao, S., Wei, Y., Buchardt, O., 2014. A PCR Based Protocol for Detecting Indel Mutations Induced by TALENs and CRISPR/Cas9 in Zebrafish. PLoS One 9, e98282.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V, Zhang, F., 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759–771.

Zhang, L., Liu, W., Shao, C., Zhang, N., Li, H., Liu, K., Dong, Z., Qi, Q., Zhao, W., Chen, S., 2014a. Cloning, expression and methylation analysis of *piwil2* in half-smooth tongue sole (*Cynoglossus semilaevis*). Mar. Genomics 18 Pt A, 45–54.

Zhang, X., Wang, H., Li, M., Cheng, Y., Jiang, D., Sun, L., Tao, W., Zhou, L., Wang, Z., Wang, D., 2014b. Isolation of Doublesex- and Mab-3-Related Transcription Factor 6 and Its Involvement in Spermatogenesis in Tilapia1. Biol. Reprod. 91, 136.

Zhang, Y., Chen, J., Cui, X., Luo, D., Xia, H., Dai, J., Zhu, Z., Hu, W., 2015a. A controllable on-off strategy for the reproductive containment of fish. Sci. Rep. 5, 7614.

Zhang, Z., Lau, S.-W., Zhang, L., Ge, W., 2015b. Disruption of Zebrafish Follicle-Stimulating Hormone Receptor (*fshr*) But Not Luteinizing Hormone Receptor (*lhcgr*) Gene by TALEN Leads to Failed Follicle Activation in Females Followed by Sexual Reversal to Males. Endocrinology 156, 3747–3762.

Zhao, H., Duan, J., Cheng, N., Nagahama, Y., 2012. Specific expression of *Olpiwi1* and *Olpiwi2* in medaka (*Oryzias latipes*) germ cells. Biochem. Biophys. Res. Commun. 418, 592–597.

Zhong, Z., Niu, P., Wang, M., Huang, G., Xu, S., Sun, Y., Xu, X., Hou, Y., Sun, X., Yan, Y., Wang, H., 2016. Targeted disruption of sp7 and myostatin with CRISPR-Cas9 results in severe bone defects and more muscular cells in common carp. Sci. Rep. 6, 22953.

Zhou, L., Feng, Y., Wang, F., Dong, X., Jiang, L., Liu, C., Zhao, Q., Li, K., 2018. Generation of all-male-like sterile zebrafish by eliminating primordial germ cells at early development. Sci. Rep. 8, 1834.

Zhou, L., Gui, J., 2017. Natural and artificial polyploids in aquaculture. Aquac. Fish. 2, 103–111.

Zhu, B., Ge, W., 2018. Genome editing in fishes and their applications. Gen. Comp. Endocrinol. 257, 3–12.

Zhu, X., Xu, Y., Yu, S., Lu, L., Ding, M., Cheng, J., Song, G., Gao, X., Yao, L., Fan, D., Meng, S., Zhang, X., Hu, S., Tian, Y., 2014. An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. Sci. Rep. 4.

Zhu, X., Ye, K., 2012. Crystal structure of Cmr2 suggests a nucleotide cyclase-related enzyme in type III CRISPR-Cas systems. FEBS Lett. 586, 939–945.

Zischewski, J., Fischer, R., Bortesi, L., 2017. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. Biotechnol. Adv. 35, 95–104.

Zohar, Y., 1989. Endocrinology and fish farming: Aspects in reproduction, growth, and smoltification. Fish Physiol. Biochem. 7, 395–405.

Zohar, Y., Gothilf, Y., Wray, S., 2007. Inducing sterility in fish by disrupting the development of the GnRH system. US7194978B2.