Generation of transgenic zebra finches by the culture and genetic modification of germline stem cells

Doctoral thesis for obtaining the

academic degree

Doctor of Natural Sciences (*Dr. rer. nat.)*

submitted by

Gessara, Ivana Lis

at the

Universität Konstanz

Faculty of Sciences

Department of Biology

Konstanz 2021

Date of the oral examination: June 25th 2021
1. Reviewer: Prof. Martin Wikelski
2. Reviewer: Prof. Manfred Gahr

Table of Contents

Summary

Genes are DNA segments that encode the instructions to make particular proteins or RNA molecules (Alberts et al., 2008). They are the units of heredity and together with the environment they shape the functioning and phenotype of all living creatures. Small alterations in the expression of a gene can have implications in the performance of complex networks connecting DNA, RNA and proteins leading to adaptation or disfunction. Therefore, experimental methods to modify gene expression are essential to reach a deeper knowledge of the gene's function. Animal species with structural, functional or developmental similarities to man are used with great success to study gene function, for example in the field of human neurodegenerative diseases. Different methods were developed to produce animal models with an alteration in the expression of a particular gene to comprehend the mechanisms behind genetic diseases and to come up with gene therapies to cure them. Songbirds in particular are commonly used as an animal model to study the genetic and neural basis of vocal learning and communication, as well as sex hormone dependent brain development and adult neurogenesis (Goldman and Nottebohm, 1983; Paton and Nottebohm, 1984; Gahr, 2007; Balthazart et al., 2010; McCarthy and Arnold, 2011; Mello, 2014; Prather et al., 2017; Mooney, 2020). Songbirds are vocal learners that possess many parallels with speech acquisition in humans and therefore constitute an excellent animal model to study speech and language disorders (Abe et al., 2015; Liu et al., 2015). In the scope of this thesis I have developed a new method to produce transgenic songbirds by the culture and genetic modification of germline stem cells (Gessara et al., 2021), since genetic manipulation in songbirds has not yet worked routinely. For this it was necessary to examine and optimize the genetic signalling pathways and biological factors of the culture of primordial germ cells. To prepare a first application of my transgenic technology, I focused on the song control system, responsible for learning and production of song in songbirds. I have studied the structure and function of physiological variations of the androgen receptor and propose a method for the genetic alteration of the androgen receptor in a zebra finch transgenic model to further investigate its role in song learning and development.

Zusammenfassung

Gene sind DNA Abschnitte die für die Herstellung bestimmter Proteine oder RNA-Moleküle (Alberts et al., 2008) kodieren. Diese Einheiten der Erbinformation bestimmen, zusammen mit Umwelteinflüssen, den Genotyp und die Lebensfunktionen aller Lebewesen. Schon kleine Veränderungen bei der Expression von Genen können sich auf das Zusammenspiel von DNA, RNA und Proteinen auswirken und so zu Anpassungen, aber auch Fehlfunktionen, im Organismus führen. Deshalb ist es nötig Genmanipulationsmethoden zu entwickeln, durch deren Einsatz man ein tieferes Verständnis der Genfunktion erreicht. Tiermodelle die strukturelle, funktionelle oder entwicklungsbiologische Übereinstimmungen mit dem Menschen aufweisen wurden mit Erfolg als Modell verwendet, um zum Beispiel neurodegenerative Erkrankungen zu erforschen. Es wurden verschiedenste Techniken etabliert um Tiermodelle zu entwickeln und einzelnen Gene zu manipulieren. Diese Genveränderungen erlauben dann Mechanismen hinter Erkrankungen zu verstehen und letztendlich Therapien zu entwickeln um diese zu heilen. Singvögel werden gerne als Tiermodell um die neuronalen Grundlage von vokalen Fähigkeiten und die darüber laufenden Kommunikation verstehen zu lernen. Des weiteren eignen sich Singvögel als Modell für geschlechtshormonabhängige Gehirnentwicklung und adulte Neurogenese (Goldman and Nottebohm, 1983; Paton and Nottebohm, 1984; Gahr, 2007; Balthazart et al., 2010; McCarthy and Arnold, 2011; Mello, 2014; Prather et al., 2017; Mooney, 2020). Singvögel müssen ihren Gesang erlernen, genau wie der Mensch Sprache erlernen muss. Beide Prozesse weisen Parallelen auf und der Gesang des Vogels ist deshalb ein hervorragendes Modell um Sprache und Sprachstörungen zu untersuchen (Abe et al., 2015; Liu et al., 2015). Im Rahmen dieser Doktorarbeit habe ich eine Methode entwickelt die es erlaubt über die Kultivierung und Manipulation primordialer Keimzellen transgene Singvögel zu erzeugen, da die Genmanipulation in Singvögeln bisher nicht routinemässig funktioniert. Dazu war es notwendig die genetischen Signalwege und biologischen Faktoren der Kultur von primordialen Keimzellen näher zu untersuchen und zu optimieren. Zum Abschluss habe ich mich auf die Funktion verschiedener physiologisch relevanter Variationen des Zebrafinken-Androgenrezeptors konzentriert und Vorschläge erarbeitet welche genetischen Veränderungen für die Erforschung des Gesangslernen besonders geeignet wären.

General Introduction

The zebra finch as an animal model

The zebra finch is part of the bird family Estrildidae (Vieillot 1817). Zebra finches are social and gregarious birds that vocalize almost constantly (Zann 1996). They are socially monogamous and form strong social bonds (Birkhead et al., 1988; Schielzeth and Bolund, 2010; Brandl et al. 2019). Their reproductive cycle is mainly related to the availability of resources, therefor being called "opportunistic breeders" (Immelmann, 1962). This behavior makes keeping and breeding zebra finches easy under laboratory conditions.

The zebra finch (Taeniopygia guttata) was initially introduced as an animal model to study plumage sexual dimorphisms (Morris, 1954; Immelmann, 1962; 1969). The males are characterized by brown flanks with white points, a narrow black chest band, an orange patch on the cheeks and an intensely orange/red beak. The females, on the other hand, are more drab. In females the orange cheek patches and brown flanks are missing, as well as the black chest band (Figure 1). This plumage dimorphism is developed during their first moult, which is about 25 days after hatching (Zann, 1996). Before that, they have uniformly grey juvenile plumage and show little gender-specific phenotype. Next to this plumage dimorphism sex specific behaviour can be observed, as males differ fundamentally from females in that they are capable of learning and producing a courtship song. Song learning and production comes associated to a complex network of brain nuclei that constitute the neural correlate for vocal learning and vocal production. Because of the sexual dimorphism of the neural song control system, the zebra finch was used over the past four decades as a major model for anatomical, molecular, endocrinological and neurobiological studies in avian species (Nottebohm and Arnold, 1976; Gurney and Konishi, 1980; Konishi and Gurney, 1982; Gahr et al., 1987; Gahr, 1990; Gahr and Metzdorf, 1997; Dittrich et al., 1999; Agate et al., 2003; Derégnaucourt et al., 2005; Aronov et al., 2008; Derégnaucourt and Gahr, 2013; Frankl-Vilches et al. 2015).

Figure 1. Zebra finch plumage sexual dimorphism. Zebra finch male (left) and female (right)

Song development in the zebra finch

The group of the songbirds (order: OSCINES), with about 4000 described species constitutes the most extensive taxon of birds (Gill and Donsker, 2014) and has developed the ability of song learning. The song learning process takes place in a so-called critical phase, which is very variable over time depending on the species (Thorpe, 1954; Konishi, 1965; Marler, 1970). In the case of the zebra finch, this section is limited in time and comprises three months (Arnold, 1975; Zann, 1996; Mooney 1999). Because of this limited learning phase of the zebra finch, it becomes part of the " age-limited learners" group. The learning phase can be divided into two time windows, which in some species such as the zebra finch partially overlap with each other (Konishi 1965). In the first period of time, which is known as the “sensory phase”, the juveniles memorise the song ("template") of one or more male conspecifics "Tutors" (Immelmann, 1969; Slater et al., 1988). This phase begins about 25 days after hatching in the zebra finch and lasts between two and five weeks (Böhner, 1990; Roper and Zann, 2006). At around 30 days, the young bird begins to sing to what has been learned from the template (Arnold, 1975; Zann, 1996), which marks the beginning of the "Sensorimotor phase". In the weeks that follow, the singing develops, being still plastic and therefore referred to as "subsong". Starting with this phase the bird adapts its motor and neural skills to adapt their song to the previously learned template. At around 100 days of age, the song learning process is complete (Zann, 1996). From then on, the zebra finch utters a very stereotypical song

pattern, which will remain constant for a long time (Immelmann, 1969) (Figure 2). In relation, between the fourth and 15th month after hatching only minor changes in the song pattern were noticed (Pytte et al., 2007; McDonald and Kirn, 2012).

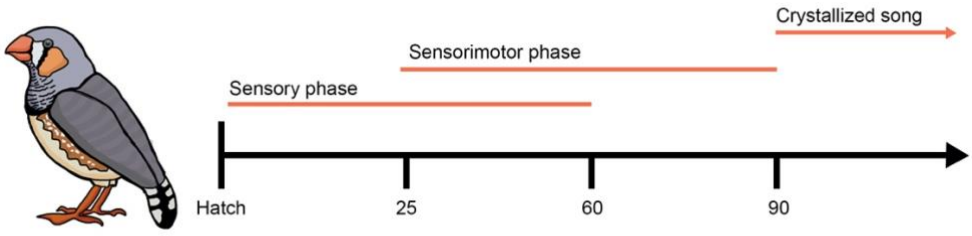


Figure 2. Development of the song in age-limited learners. Age-limited learners such as the zebra finch will memorize the tutor song in the so-called sensory phase. They will then practice the song in the overlapping sensorymotor phase until their song becomes crystalized. After this happens their song will not have significant further modifications.

The song control system, the neural base for singing

In order to be able to hear and process the tutor song and to learn and produce a copy of it, songbirds have evolved complex telencephalic nuclei dedicated to song perception, learning and production. Together they form the so-called song control system (Nottebohm et al. 1976; Nottebohm, et al., 1982; Fortune and Margoliash 1995; Bolhuis and Gahr 2006). The song control system receives the auditory input and includes projections to the motor neurons responsible for the control of the respiratory muscles and the syrinx, the vocal organ in birds. The song control system is subdivided in two pathways, the song motor pathway (SMP) and the anterior forebrain pathway (AFP) essential for song production and song learning respectively. The song motor pathway is formed by two song nuclei that project to motor neurons of the respiratory system and syrinx, HVC (a letter based name formerly known as high vocal centre) and RA (robust nucleus of the arcopallium) (Wild, 2004). The anterior forebrain pathway is formed by a pallial-basal ganglia-thalamic loop that includes the projection from HVC to the basal

ganglia nucleus Area X. Area X projects to the thalamic nucleus DLM (medial nucleus of the dorsolateral thalamus), which projects to pallial nucleus lMAN (lateral magnocellular nucleus of the anterior nidopallium). LMAN closes the loop back to Area X and also projects to RA (Doupe et al., 2004; Nottebohm et al., 1976). Each of these nuclei plays a role in song learning and/or production (Figure3). These neuronal accumulations (the song control areas) differ from the surrounding parenchyma by electrophysiological, molecular and genetic parameters (Nottebohm et al., 1976; Margoliash, 1997; Wild, 1997; Warren et al., 2010; Lovell et al., 2018) and have numerous functional and developmental similarities with areas of the mammalian neocortex (Reiner et al., 2004), which made songbirds an ideal animal model to study speech and vocal impairments in humans as well as song learning and song production in birds (Nottebohm et al., 1976; Fortune and Margoliash 1995; Hahnloser et al., 2002; Thompson et al., 2007; Sebastian Haesler et al., 2007; Aronov et al., 2008; Bolhuis, Fee and Scharff 2010; Bolhuis et al., 2010; Liu et al. 2015; Konopka and Roberts 2016; Nieder and Mooney 2020).

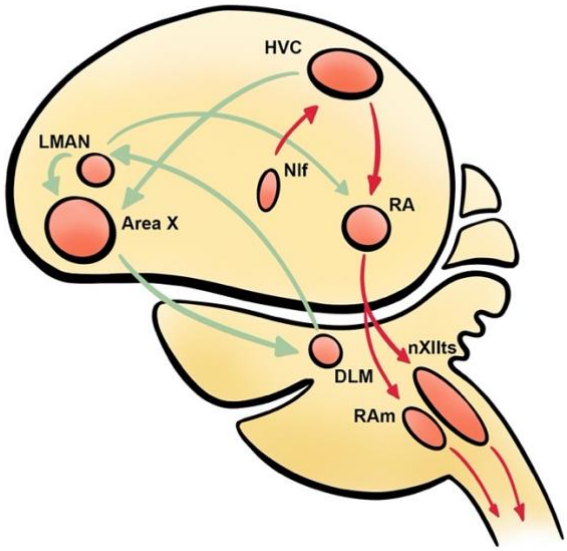


Figure 3. The song control system. Representation of the song control system in the songbirds brain. The song control system is divided in the Song Motor Pathway (SMP, red arrows) and the Anterior Forebrain Pathway (AFP, green arrows). The SMP connects HVC and RA with motor neurons innervating the syrinx and the respiratory system. The

AFP conforms a loop connecting HVC with Area X, area X with DLM, DLM with lMAN and lMAN to RA as well as back to Area X.

Sexual dimorphism in the brain

The song of the zebra finch is a classic example of sexually dimorphic behaviour. In many songbird species, the males are commonly the ones producing complex learned songs while only a low percentage of females sing sporadically (Riebel et al., 2019; Ko et al., 2020). This sexual dimorphism it is also present in the brain, where the song control system develops differently in males and females. Not only the size of the song nuclei differs being larger in males than in females (Nottebohm and Arnold, 1976; Konishi and Akutagawa, 1987; Konishi and Akutagawa, 1990; Adret and Margoliash, 2002; Wade and Arnold, 2004) but also the interconnections between nuclei are more prominent in males (Nordeen et al., 1987; Nordeen and Nordeen, 1989; Mooney and Rao, 1994) and the neuronal soma, number of dendrites and density of synapses are larger in the song control system of males than in females (DeVoogd and Nottebohm; 1981; Nealen, 2005).

Up to the age of ten days, the HVC of males and females does not differ in relation to the number and volume of neurons (Kirn and DeVoogd, 1989; Nixdorf - Bergweiler, 1996), but differences were found in regard to the number of the androgen receptor - expressing cells (Gahr and Metzdorf, 1999). In the male bird, the HVC then grows linearly with the telencephalon until the 20th post hatching day (PHD). After that HVC starts growing more than the surrounding area reaching its adult volume around the 60th PHD (Bottjer et al., 1985; Nixdorf-Bergweiler, 1996). In the female zebra finch, however, the HVC stops expanding and/or degenerates between the tenth and 40th PHD, due to apoptotic processes (Kirn and DeVoogd, 1989; Gahr and Metzdorf, 1999) and in the adult female it only reaches approx. 16% of the male volume (Nixdorf-Bergweiler, 1996; Gahr and Metzdorf, 1999).

Genetic and hormonal basis of sexual dimorphism

Sex differences arise in response to diverse sex-specific signals originating from inherent differences in the genome and involve cellular mechanisms that are specific to individual

tissues or brain regions (McCarthy and Arnold 2011). The sexually dimorphic brain development in mammals and birds is determined by genetic factors (Carrer and Cambiasso, 2002; Agate et al., 2003; 2004; Gahr, 2003), that lead to the expression of sex hormone receptors such as the androgen (AR) and estrogen receptor (ER) (McCarthy, 2016). After binding of their ligands sex hormone receptors are activated, enter the nucleus and induce gene expression cascades triggering the differential brain development (Balthazart et al., 1992; Gahr et al., 1993; Jacobs et al., 1996; Gahr, 2007; Frankl-Vilches and Gahr, 2017). In this context the androgen receptor (AR) is found expressed in the song control system in HVC, RA, and lMAN of all songbirds studied (Balthazart et al. 1992; Bernard et al. 1999; Gahr et al. 1998, 2008; Metzdorf et al. 1999; Fusani et al. 2000; Voigt and Gahr 2011; Fraley et al. 2010; Quispe et al. 2016) as well as in mMAN (medial magnocellular nucleus of the anterior nidopallium) and NIF (nucleus interfacialis) in canaries and zebra finches (Balthazart et al. 1992; Metzdorf et al. 1999; Fusani et al. 2000). Androgen receptors were also found in Area X of zebra finches and canaries but only in some individuals for unknown reasons (Gahr 2004; Kim et al. 2004). In contrast to the widespread expression of the androgen receptor in the song control system, the expression of the estrogen receptor is limited to the caudomedial HVC in zebra finches and throughout HVC in canaries and an area dorsally from RA (Figure 4) (Gahr et al., 1987; Gahr and Konishi, 1988; Gahr and Metzdorf, 1997; 1999; Metzdorf et al., 1999; Gahr, 2004; Kim et al., 2004; Peterson et al., 2005). Since the same vocal control areas contain androgen and estrogen receptors in male and female songbirds (Gahr and Konishi 1988; Gahr et al. 1993, 1996; Metzdorf et al. 1999; Gahr and Metzdorf 1999; Jacobs et al. 1999; Kim et al. 2004) sex hormones can affect the vocal control areas in both males and females. In order to have a better understanding of how androgen and estrogen receptors affect the development of the song control system and the vocal learning process in songbirds experimental approaches are necessary to alter their expression.

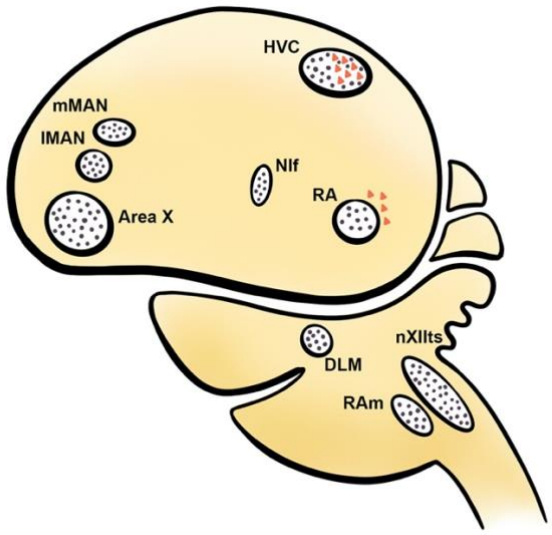


Figure 4. Androgen and estrogen receptor distribution in the song control system. Representation of the songbird brain with the nuclei of the song control system that contain cells expressing the estrogen receptor (orange triangles) in HVC and an area dorsally to RA and the androgen receptor (purple dots) in HVC, mMAN, lMAN, Area X, Nlf, RA, DLM, Ram and nXIIts.

Non-transgenic methods used to alter gene expression in birds

Different methods were developed in avian species in order to determine the function of particular genes and their participation in developmental and physiological processes. The most widely used non transgenic methods include inhibitory drugs, viral vectors and *in vivo* electroporation.

Inhibitory drugs

To determine the function of genes in a systemic way, other methods were developed. Inhibitor drugs can be used to inactivate or antagonize the effect of a particular compound to determine their role in a particular process. These substances were most often applied systemically but can also administered locally. Fadrozole, that inhibits the metabolization of testosterone into estrogen, was injected systemically or locally in the brain of zebra finches to study its role in the brain development and activation of avian sexual

behaviours (Wade et al., 1994; Alward et al., 2016). Similarly, the estrogen receptor antagonist ICI 182,780 was injected intracranially in young zebra finches to investigate the role of estrogens in the dimorphic brain development (Bender et al., 2008). Flutamide, an androgen-receptor blocker was incorporated systemically in developing zebra finches to investigate the role of androgens in the sexual differentiation of the zebra finch song system and in the song production (Bottjer et al., 1992; Grisham et al., 2007). Inhibitors have the benefit that their effect can be modulate in a doses basis but it is difficult to design and test the specificity of a drug in its function toward a particular gene product.

Viral vectors

Viral vectors were developed to deliver genetic material and alter gene expression in a more specific way. Lentiviral (LV) and adeno-associated viral (AAV) vectors can be designed to alter the expression of a particular gene after local injection and transduction of their target cells. In avian species viral vectors can be used for anatomy and circuit function in the brain (Düring et al., 2020) as well as for the study of genes involved in human speech impairment diseases (Haesler et al., 2007; Adam et al., 2016; Norton et al., 2019). This method is ideal for local genetic alterations. However, the disadvantage is that it is necessary to know in advance where and when the gene might play a role. Further, in the case of LV possible insertion into genes might lead to disfunction which is avoided when using AAV due to their transient effect.

In vivo electroporation

DNA can be delivered into the neural tube of chicken embryos or the brain of early hatched zebra finches by electroporation. Transfection of an expression plasmid together with injection of transposase in the case of transposon electroporation is a technique that has been introduced for genetic alteration in chicken embryos (Itasaki et al., 1999) and recently applied to zebra finch embryos (Ahmadiantehrani and London, 2017). The disadvantages are that there is no control over cell specificity or localization specificity and so far it is only applicable to embryos.

Methods developed for the generation of transgenic animal models

Transgenic animal models consist of genetically altered animals where the expression of a gene was altered or a foreign gene was introduced into the genome. Transgenic animals have the ability to pass the transgene to their offspring offering the possibility to obtain a large number of identically modified specimens. Initially transgenic animals were produced to study the effects of gene expression changes in the whole animal. More recently, animal models are used where the effects are limited to specific organs or brain regions by crossing transgenic animals that express the Cre-recombinase under the control of a region specific promoter with a transgenic animal where the genetically changed gene sequence is flanked by two loxP sites. Further, transgenic models can be developed with gene switch mechanisms inducible by light or antibiotics (Rao et al., 2009; Madisen et al., 2012).

Development of transgenic technologies

Genetically modified animals have been of a big help in many research fields such as life science, biomedical engineering and pharmacology allowing researchers to treat diseases and to manipulate gene expression in order to get more insights on gene expression, function and regulation. In 1973 Rudolf Jaenisch injected a simian virus into a mouse embryonic blastocoel cavity creating the world's first transgenic animal (Jaenisch and Mintzt 1974), yet these mice were not capable of passing the transgene to the offspring. In 1980 Frank Ruddle injected foreign DNA into the pronucleus of a mouse zygote, which was incorporated into the genome of the host, giving bird to animals carrying the transgene that were able to pass it to their offspring (Gordon et al. 1980).

Gene targeting in embryonic stem cells

Gene targeting techniques in mouse stem cells were first developed in the 1980s, when they were used to generate knock out transgenic mice models where a gene was ubiquitously inactivated (Thomas and Capecchi, 1987; Thompson et al., 1989). Mouse stem cells can be extracted from an early embryo and expanded *in vitro* conserving their

ability to differentiate into different cell types. To generate targeted mutagenesis initially cultured stem cells were initially genetically modified by the insertion of vectors containing the desired mutation. These vectors paired with the cognate chromosomal DNA sequence and transferred the mutation to the genome by homologous recombination. The targeted stem cells were then reintroduced into morula or blastocyst stage embryos generating chimeras formed of cells derived from both the donor and the host cells (Capecchi 1989). Alternatively, genes could also be disrupted by transpositional 'gene-trapping' mutagenesis. Gene trapping consist in the incorporation of a construct containing a reporter gene that will be inserted into the target gene and will disrupt its sequence inactivating its expression and expressing the reporter gene instead (Evans et al., 1997).

Pronuclear injection-based transgenesis

To produce transgenic animals with this method a DNA construct is microinjected into the pronucleus of a one-cell-stage zygotes (Bockamp et al., 2008). The DNA sequence will then insert randomly at one or more genomic loci as either a single or as multiple copies. When the embryonic cells start dividing all the daughter cells will have at least one copy of the transgene. This method is limited by the uncertainty of the level of gene expression due to the random insertion site and copy number. Recently, with the development of new gene targeting technologies the integration of a single-copy transgene in a specific location in the genome, changes of the DNA-sequence or epigenetic modification became possible.

Viral transfection for transgenesis

Lentiviruses constitute an alternative to the use of embryonic stem cells or pronucleus injections of DNA vectors to generate transgenic animal models. The lentiviral vectors are produced replacing non-essential viral genes with foreign genes of therapeutic interest and programmable endonucleases such as TALENs or CRISPR/Cas9. A viral vector includes an expression cassette for gRNA or for the transgene followed by a woodchuck post regulatory element (WPRE) to increase the transcript stability (reviewed by Kay et al., 2001 Pfeifer 2004). In lentiviral vectors long terminal repeats (LTR) 5' and 3' of the

expression cassette have four functions: to act as an RNA polymerase II promoter, to terminate transcription, to promote polyadenylation and to act as recognition site for insertion in the host genome. Recombinant viral vectors are injected into the pronucleus, reverse transcribed and integrated in the genome, altering the gene expression of the embryo (Pfeifer et al., 2002; Lois et al., 2002). Over the past two decades transgenic animal models of mouse, rat, cows and pigs among other animals have been developed using lentiviral vectors with high transgenic efficiencies (Ikawa et al., 2003; Michalkiewicz et al., 2007; Whitelaw et al., 2004; Ewerling et al., 2006).

Sperm mediated gene transfer and testis mediated gene transfer

Sperm mediated gene transfer consists in the introduction of foreign genes into sperm cells *in vitro* followed by an artificial insemination of the female with the genetically modified sperm (Lavitrano et al., 1989, Maione 1998, Smith and Spadafora 2005). In contrast to this, testis mediated gene transfer, consists in the *in vivo* injection of DNA into the testis followed by testicular electroporation to target and modify germ cells and produce genetically altered sperm (Huang et al., 2000; Dhup and Majumdar 2008).

Modern gene editing techniques

Having access to the one cell embryo opened the possibility to the application of many different gene editing techniques that result convenient for different transgenic animal models. In the following I will introduce some of the most commonly used and more advanced gene editing techniques. I will discuss whether they can be applied to produce transgenic songbirds or new methods need to be develop in other to enable their use.

Conditional mutagenesis

The necessity of the development of conditional mutations emerged when studying genes in which a homozygous mutation would be lethal. Conditional mutations consist of mutations that are time- (endogenously controlled; experimentally inducible) or tissue-specific. To use this method palindromic recombinase target sequences are used to flank a

specific region of the target gene. This sequences commonly include the Cre-associated loxP sequence or the Flp-associated FRT sequence (Bouabe and Okkenhaug, 2013). When these sequences are exposed to the associated recombinase enzyme they recombine with each other to remove or invert a DNA sequence from the target gene inactivating its expression (Figure 5). Associated recombinase enzymes can be incorporated in the genome with a responsible element to a controlable transactivator protein (CTP) to be able to control its expression. The CTP gene is inserted in the genome preceded by an ubiquitously expressed promoter or a cell/tissue specific promoter (for a cell/tissue specific expression). The CTP gene expression can be activated by an antibiotic such as doxycycline or tetracycline which will stimulate the production of the CTP that will then bind to its responsible element and activate the expression of the associated recombinase enzyme gene. The associated recombinase enzyme will then be produced and will act on the LoxP sites generating a time and tissue conditional mutation (Figure 5) (Gossen and Bujard, 1992; Kistner et al., 1996; McLellan et al., 2017).

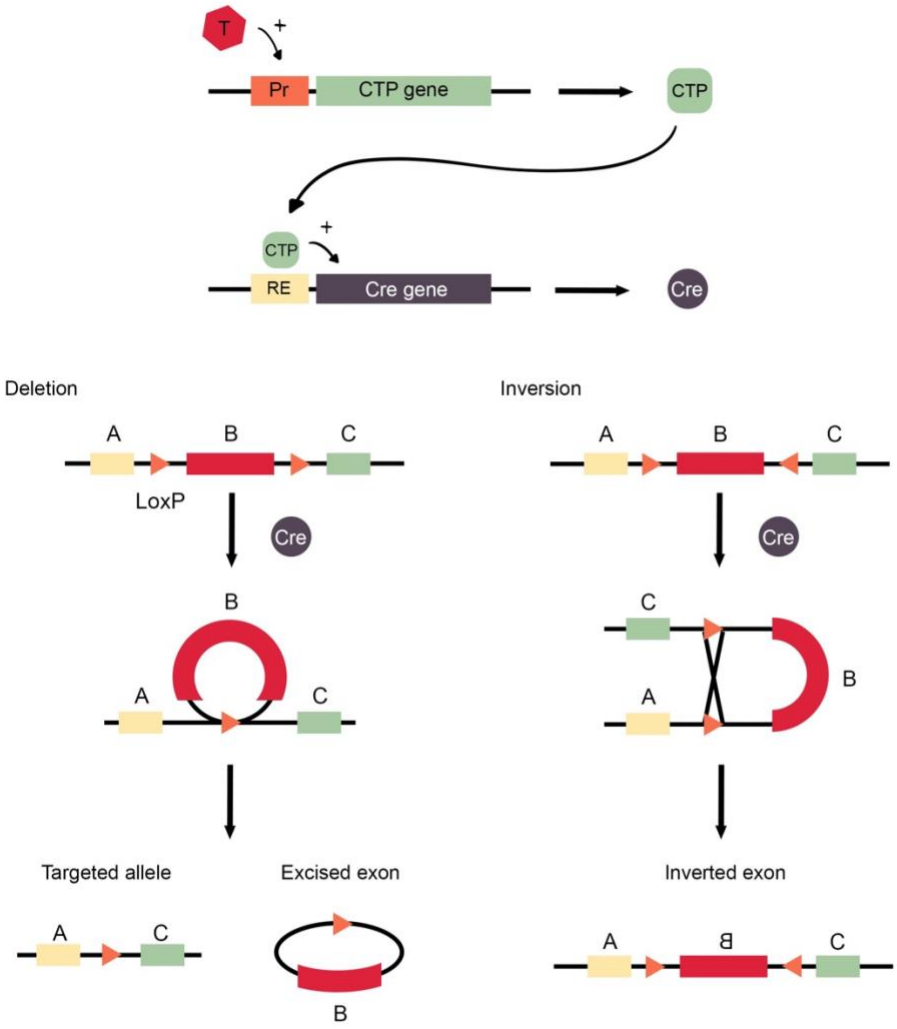


Figure 5. Conditional mutagenesis. In this example the incorporation of the antibiotic tetracycline (T) activates the expression of the controllable transactivator protein gene (CTP gene) which will start the synthesis of the controllable transactivator protein (CTP). The CTP will then interact with the responsible element (RE) in front of the Cre gene and activate its expression. The Cre protein will then be synthesised. After Cre protein expression two effects of Cre-recombinase mediated conditional mutagenesis are represented: Deletion when a loop is formed between unidirectional loxP site; Inversion when the loxP sites are oriented toward each other.

Knock-down gene expression using RNAi

A knock-down leads to a partial reduction of the expression of a gene. It is an alternative to conditional mutagenesis for genes that can't be completely inactivated due to its lethal

consequences. The generation of transgenic knock-down animals by the use of RNAi consists in the introduction of a double stranded RNA (dsRNA) or a short hairpin RNA (shRNA) that is homologous a mRNA sequence and will result in the posttranscriptional silencing of the gene (Fire et al., 1998). RNase III family nucleases (DICER) cut the expressed ds RNA or shRNA into small fragments the so-called small interfering RNAs (siRNAs). SiRNAs are then incorporated into the RISC (RNA induced silencing complex) that uses the siRNA sequence as a guide to identify the homologous mRNAs which is subsequently cleaved (Figure 6). This method was used to generate transgenic animals through transduction of lentiviral vectors expressing siRNA into preimplantation mouse embryos (Tiscornia et al., 2003) or pronuclear injection of constructs that express shRNAs (Chang et al., 2004; Seibler et al., 2007; Dickins et al., 2007). The disadvantage of siRNA is the low specificity with off-target effects (Jackson and Linsley 2010).

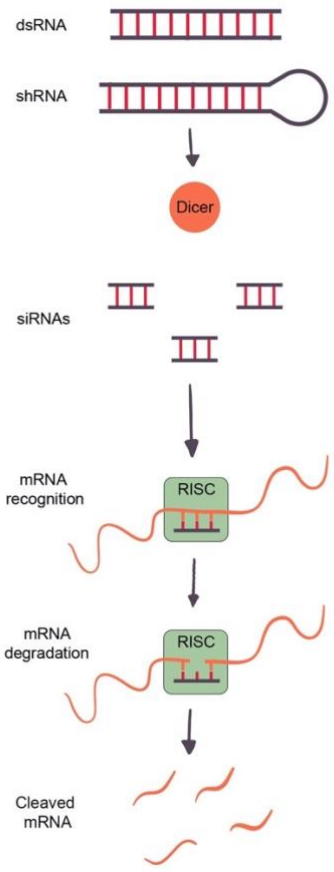


Figure 6. The mechanism of a RNAi knock-down. Double stranded RNA (dsRNA) or short hairpin RNA (shRNA) under control of a RNA-polymerase III promoter are delivered into the cells by a vector. Once expressed they are cut into small interfering fragments (siRNA) by a DICER nuclease. The siRNA is then incorporated into the RNA

induced silencing complex (RISC) where it will be used as a guide to cleave the targeted mRNA.

Knock-down of gene expression using dominant negative variants.

Dominant negative gene variants contain deletions that inactivate one or more functions of the encoded protein. Once this genetically modified protein is produced by the cells it is functionally competing with the native variant. The effects of the genetically modified variant can be dominant when its expression level is higher than that of the native version. Thus, binding of a ligand to a dominant negative receptor variant might not generate the subsequent cascade of signalling events and lead to a reduction of the ligand effects (Herskowitz 1987) (Figure 7). Dominant negative variants can be incorporated in the genome after pronuclear injection or viral transduction (Amendt et al., 1998; Jones et al., 1999).

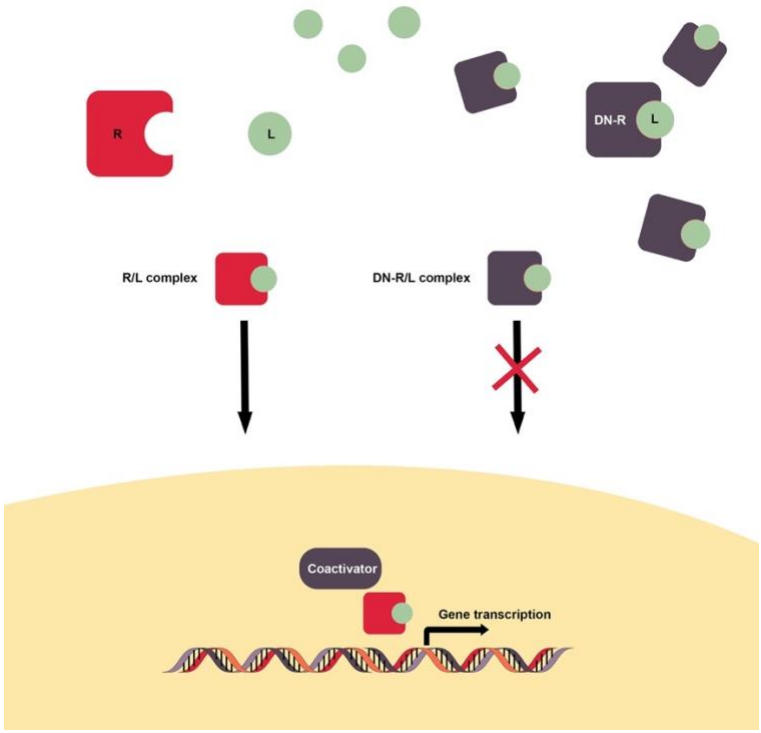


Figure 7. The mechanism of a knock-down using a dominant negative gene variant. As an example a dominant negative variant of a receptor (DN-R) is shown that lacks part

of the DNA sequence responsible for its transactivational effect. Once this gene is expressed under control of a highly active promoter, the DN-R may outcompete the native receptor (R). The DN-R can still bind to its ligand (L) but will not activate gene transcription reducing the effect of the native variant.

Programmable endonucleases for genome editing

Programmable endonucleases can produce double cuts in a gene and introduce genetic mutations in two ways: by non-homologous end joining (NHEJ) an error prone reparation process that leads to mutations (Indels) that can inactivate the gene, or by homology-directed repair (HDR) when the cell is provided with a DNA template that will be used as a reference to repair the gene (Joung and Sander, 2013; Gaj et al., 2013; Sander and Joung, 2014).

There are four methods developed for gene editing that use programmable endonucleases: and these are homing endonucleases (HEs); zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs); and the clustered regularly interspaced short palindromic repeats associated 9 (CRISPR/Cas9) system. Initially, HEs were used to increase gene-targeting efficiency in ES cells (Smih et al., 1995). Later ZFNs were developed and used for the generation of transgenic animals (Geurts et al. 2009; Meyer et al. 2010, Carbery et al., 2010). TALENs were developed in 2010, easier to produce and use than HEs and ZFNs and also used to generate transgenic models (Carlson et al. 2012; Tesson et al. 2011, Sung et al., 2013).

In late 2012 and early 2013 CRISPR/Cas9 was developed consisting in a Cas9 nuclease and a guide RNA (gRNA) (Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013). The gRNA guides Cas9 to a complementary site in the genome. When this site is located three to four base pairs upstream of the Protospacer Adjacent Motif (PAM) the Cas9 nuclease recognises the target site and produces a double strand (ds) DNA break. The ds DNA break will be then repaired by NHEJ or HDR (Figure 8) (Jinek et al., 2012, 2013; Mali et al., 2013; Cong et al., 2013; Cho et al., 2013). The CRISPR/Cas9 system was rapidly developed into a highly efficient and fast method to produce transgenic animal models (Shen et al., 2013; Wang et al., 2013). Since then the CRISPR/Cas9 method has almost

completely overcome all other technologies for genome editing. The development of the CRISPR/Cas9 system was such a significant event in the history of gene editing that Emmanuelle Charpentier and Jennifer Doudna received in 2020 the Nobel Prize in Chemistry for its discovery.

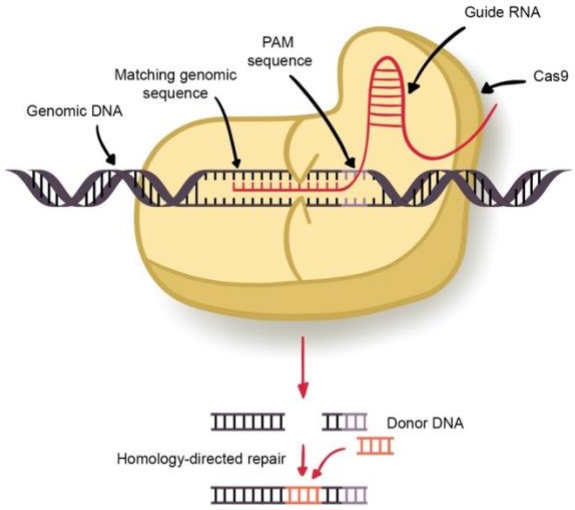


Figure 8. Mechanism of Crispr/Cas9 mediated genome editing. Once the Cas9 enzyme and the guide RNA are incorporated into the cells a double strand cut is produced in the targeted DNA three to four nucleotides upstream the PAM sequence, where it will produce a double strand cut in the DNA. The DNA cut will be repaired by non-homologous end joining (NHEJ) or homologous direct repaired (HDR) using a DNA template.

Methods developed for the generation of transgenic models in birds

Pronucleus injection, the most common technique used to generate transgenic animals, is difficult tin birds because of the opaque cytoplasm and yolk content of the avian egg. Since at the time of oviposition the avian egg contains an embryo of more than 20000 cells further limitations are imposed on the available techniques (Lee et al., 2020). Therefore, other approaches have been considered to produce transgenic birds (Fig. 9).

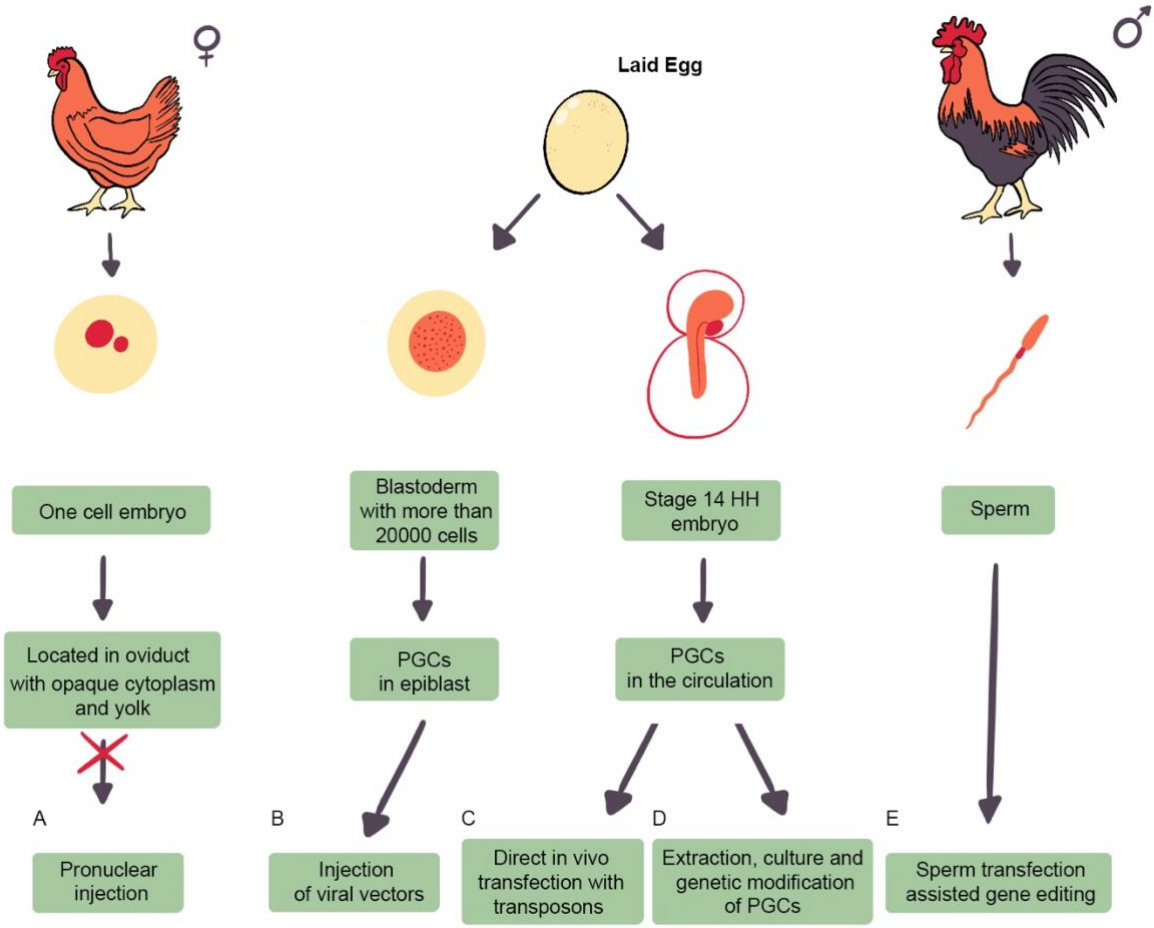


Figure 9. Methods to produce transgenic avian models. The pronucleus injection is difficult in birds due to the opaque cytoplasm and high yolk content of the one cell embryo (A). *In ovo*, Primordial Germ Cells (PGCs), the precursors of sperms and oocytes, can be genetically modified by the injection of viral vectors into the freshly laid egg (B) or by the application of transposons (C). Also, PGCs can be extracted, cultured and modified *in vitro* (D) allowing for the use of modern gene editing techniques. Finally, sperm samples can be extracted from roosters, genetically modified *in vitro* (E) by the use of modern gene editing techniques and inseminating into females. This method is difficult to apply in smaller avian species for the large sperm volume that is required. All PGC-based transgenic techniques available in birds lead to a founder generation with mosaic representation of the transgene in the gonads. Analysis of the somatic expression of the transgene requires the production of the next generation. Only in the case of Sperm Transfection Assisted Gene Editing (see below), the transgene should be expressed somatically in the immediate offspring. With all techniques the genotype of the first transgenic generation will be heterozygous.

Generation of transgenic chicken by sperm transfection assisted gene editing (STAGE)

Recently, a method was established for the generation of transgenic chicken that consists of the extraction of sperm samples followed by a transfection with Cas9-mRNA and guide RNA (gRNA) by the means of Lipofectamine (Figure 9 E). Hens were inseminated multiple times with the transfected sperm to obtain transgenic offspring (Cooper et al., 2017). Thus, a sperm volume larger than one that could be obtained from smaller bird species like the zebra finch were necessary for this procedure. An advantage of this method is that it only requires only one generation to produce transgenic animals. However, the efficiency of transgenesis was rather low 1.2% for DMRT1 and 14% for GFP in average and the edits observed in the genome were found between 50 and 200 base pairs away from the predicted CRISPR target sites (Cooper et al., 2017).

Generation of chimeras by the use of embryonic stem cells (ES)

Stem cells constitute a very efficient tool to produce transgenic mammals and it was considered to be a potentially robust method for its application to avian species. The development of a transgenic avian model by the use of stem cells was tested in the 90's. At Eyal-Giladi and Kochav stage X (EGK) (Eyal-Giladi et al., 1976; Kochav et al., 1980) the chicken embryo consists of two distinct regions, the area opaca which will develop into the extra-embryonic tissue and the area pellucida which will develop into the embryo proper (Figure 9 A). The cells present in the area pellucida are pluripotent stem cells and have been used to create chimeric birds. Stem cells from the embryos of Barred Plymouth Rock (BPR) chickens (black feathers) were injected into embryos of Dwarf White Leg Horn (white feathers) at EGK stage X, and somatic chimeras were identified by the presence of both black and white feathers (Petitte et al. 1990). Later, γ-irradiation of recipient embryos was later used to compromise the cells of the recipient embryo and improve the efficiency of transmission of the donor phenotype into the chimeras obtaining up to 85% donor cells (Carsience et al. 1993; Kagami et al. 1995; Kino et al. 1997). The injection of chicken embryonic stem cells showed no evidence of germline

transmission, contribing to somatic tissue but not to the germline which limited their use for the production of transgenic birds.

Primordial germ cells (PGCs)

Primordial germ cells are the precursors of spermatocytes and oocytes and can be found in the developing embryos. In birds PGCs arise from the epiblast and can be first detected in the central disc of the area pellucida of EGK stage X EGK embryos (Figure 10 A). At EGK stage XII the embryo is formed by two cell layers, the epiblast and the hypoblast and two cavities, the blastocoel between the epiblast and the hypoblast and the subgerminal cavity underneath the hypoblast (Figure 10 B). At Hamburger and Hamilton (HH) stage 4 (Hamburger and Hamilton, 1951) PGCs translocate to the germinal crescent, increase in number, enter the developing vascular system at HH stage 8 (Figure 10 C) and circulate in the embryonic blood stream until HH stage 14 (Figure 10 D). Finally, from HH stage 14 to 17 the PGCs leave the vascular system to migrate to their final location in the germinal ridges where they will differentiate into spermatozoa or oocytes (Figure 10 D). Methods were developed to target the PGCs in order to alter the germ line genome and produce fully transgenic birds in the next generation.

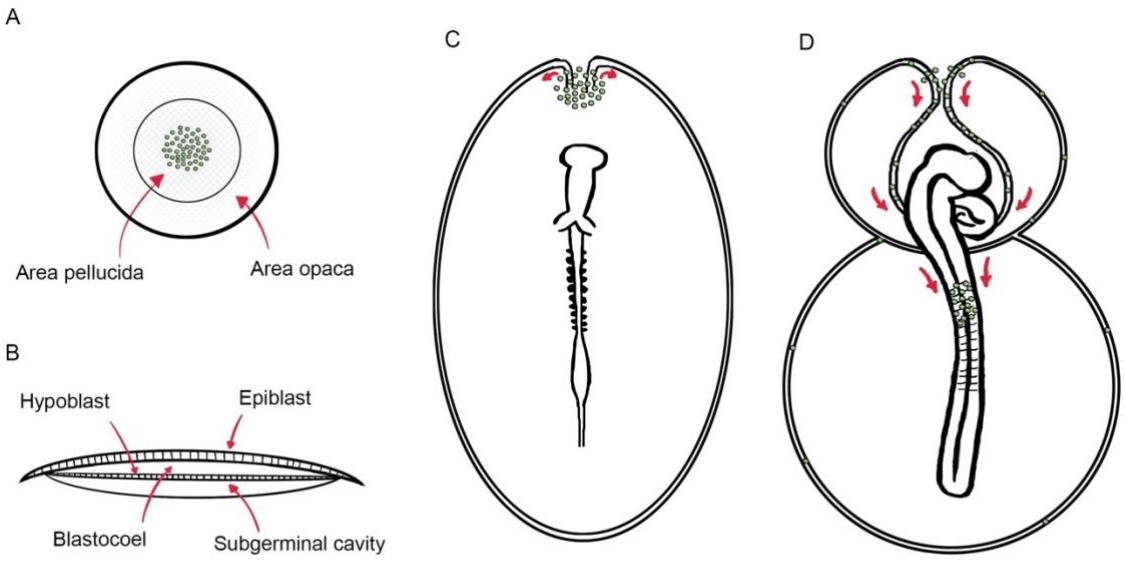


Figure 10. Avian early embryonic development and PGC migration. (A) Embryo at Eyal-Giladi and Kochav (EGK) stage X with PGCs (green) located in the centre. (B) At EGK stage XII, the embryo is formed by the epiblast, hypoblast and two cavities, the blastocoel and the subgerminal cavity. At this stage PGCs are located in the epiblast. (C) At Hamburger and Hamilton (HH) stage 10, PGCs enter the circulatory system. (D) At HH stage 14 PGCs start migrating to the gonadal anlage.

Genetic modification of PGCs by injection of viral vectors into the blastoderm

Viral vectors constitute a suitable strategy to generate transgenic birds. The first transgenic bird was generated by the injection of an avian leukosis virus into the yolk sac near to the blastoderm that successfully integrated to the germline (Salter et al., 1987). Transgenic chickens, quails, zebra finches and canaries were generated by the injection of lentiviral vectors carrying a transgene into stage X EGK or 14 HH embryos both with the aim of targeting and modifying PGCs that would differentiate into functional gametes and generate offspring with the altered genotype (Mizuarai et al. 2001; Mozdziak et al., 2003; McGrew et al. 2004; Zhang et al. 2012; Agate et al., 2009; Liu et al., 2015; Abe et al., 2015). This is so far the only method available so far for the development of transgenic songbirds and other bird species apart from chickens (Figure 9 B). However the germline transmission rated is very variable depending on the being 18% in average in chicken (Mozdziak et al., 2003; McGrew et al. 2004) 1.7% and 13% in quails depending on the injection method (Zhang et al. 2012) and 10% zebra finch but only with 3 of 23 founders successfully producing transgenic offspring (Agate et al., 2009).

Genetic modification of PGCs with transposons in-ovo

Another method developed to produce transgenic chickens consist in the transfection of PGCs *in ovo* with transposons. A transposon system is formed by two components: the transposon, a DNA sequence that is able to move from one locus in the genome to another, and the transposase, the enzyme that cleaves the DNA, releasing the transposable element. Genetically modifies chickens were generated by lipofectamine based

transfection of Tol2 transposons to target circulating PGCs at HH stage 14 (Tyack et al., 2013). Again, these methods had a low germline transmission rate 1.5% (Figure 9 C)

Generation of transgenic birds by the culture and genetic modification of PGCs

Unlike to what happens in mammals, where the PGCs reach the gonads migrating through the allantois and yolk sac, in avian species PGCs migrate to the gonads through the developing blood system (Han et al., 1994). This makes them accessible for a short period of time to be extracted from the blood, cultured and genetically modified *in vitro*. The ability to culture chicken PGCs for a long period of time opened the possibility for the application of modern genome editing techniques such as TALENs and CRISPR/Cas9 selecting targeted cells and expanding them clonally *in vitro* in this avian species. Genetically modified PGCs are then injected into EGK stage X or HH stage 14 embryos where they migrate to the gonads contribute to the germline and generate transgenic offspring (Van De Lavoir et al. 2006; Macdonald et al. 2010; Schusser et al., 2013, 2016; Dimitrov et al. 2016). The genetic modification of PGCs *in vitro* followed by injection into recipient embryos is the most effective way to produce transgenic chickens so far with a germline transmission rate of around 20% but reaching up to 95% in some particular cases (Dimitrov et al. 2016, Figure 9 D). Until now, PGCs could be expanded *in vitro* in chickens only limiting the application of this method to other bird species.

Because genetically modified PGCs compete for the production of offspring with endogenous PGCs when injected into the host embryo, therefor the germline transmission rate depends highly on the number of PGCs that were originally injected. In order to improve the germline transmission rate and thereby the number of host embryos needed, chemical and physical sterilization methods were developed to sterilized the surrogate host. Embryos were partially sterilized by injecting Busulfan into the yolk of fertile eggs (Nakamura et al., 2008) or irradiating the eggs with gamma rays (Trefil et al., 2006). In a variation of this method, genetically modified PGCs were injected into the testes of adult gamma-irradiated rosters (Trefil et al.,2017).

Transgenesis in songbirds

Until now there was only one method stablished to produce transgenic songbirds consisting on the injection of lentiviral vectors into the blastodisc of freshly laid eggs. While the time required for the generation of transgenic songbirds was short compared to methods based on cultured PGCs, this technique resulted in a germline transmission rate of 10% with only 3 of 23 founders successfully producing transgenic offspring (Agate et al., 2009). This low germline transmission is because the virus is injected into an embryo that already contains more than 20000 cells so that the odds of the virus being integrated in the PGCs population are rather low. Furthermore, this method does not allow for the application of genome editing techniques that involve the selection and clonal expansion of targeted cells. In consequence, the method restricts the application of genetic modifications to transgene over-expression, knock-down by siRNA or knock-out.

So far, over the last 15 years, three transgenic zebra finch models were developed using lentiviral vector injection into the blastoderm of songbirds. Agate and colleagues applied this method for the first time to produce transgenic zebra finches expressing enhanced Green Fluorescent Protein (eGFP) (Agate et al., 2009). In 2015, Liu and collaborators developed the first functional transgenic songbird, inserting the mutated human huntingtin gene in the zebra finch genome. This gene produces a protein responsible for the progressive deterioration of motor and cognitive function in Huntington's disease. The mutant songbirds had severe vocal disorders, including poor vocal imitation, stuttering, and progressive syntax and syllable degradation. Liu and colleagues presented an example of how a song bird can be the ideal animal model to study speech and vocal impairments in humans (Liu et al., 2015). Lastly, in 2015, Abe and colleagues developed a transgenic zebra finch model with a supressed or enhanced activity of the cAMP response element-binding protein (CREB) transcription factor. The transgenic birds showed reduced vocal learning and impaired audio-memory formation. These results demonstrate that appropriate activity of CREB is necessary for the postnatal acquisition of learned behaviour in songbirds, proving transgenic models to be an excellent method to study in more detail the mechanisms behind song acquisition and production in birds (Abe et al., 2015).

The development of a novel method to produce transgenic songbirds by the means of the extraction, culture and genetic modification of blood-borne zebra finch primordial germ cells (zfPGCs) would open the possibility for the application of targeted gene editing techniques on zfPGCs. Genetically modified PGCs could later be injected in recipient

embryos where they would migrate to the gonads and contribute to the germline generating transgenic offspring in the following generation. With this method a larger germline transmission rate than that obtained with the previously developed method could be achieved due to the injection of a large number of PGCs to colonize the gonads. The goal of my PhD was to develop and implement such method, stablishing the right culture conditions to expand zfPGCs *in vitro*, altering their genome and injecting them into recipient embryos to generate transgenic offspring.

The androgen receptor

Songbirds are major animal models to study the genetic and neural basis of vocal learning and communication (Mooney, 2020; Mello, 2014; Prather et al., 2017) as well as sex hormone dependent brain development (Gahr, 2007; Balthazart et al., 2010; McCarthy and Arnold, 2011) and adult neurogenesis (Goldman and Nottebohm, 1983; Paton and Nottebohm, 1984). Zebra finch transgenic models constitute a big step towards the understanding of these processes allowing for the study of the participating genes in a molecular and functional way.

Sex hormones and their role in brain development

Sex hormones play an essential role in brain development in mammals and birds. Studies y rodents showed that testosterone in particular stimulates cell survival (Hamson et al., 2013; Spritzer and Galea, 2007) and synaptogenesis in both males and females (Leranth et al., 2003; Lewis et al., 1995). It also regulates amygdala volume (Cooke et al., 1999), dendritic length and soma size of adult motor neurons (Kurz et al., 1986), and the density of dendritic spines in the male preoptic area, a brain area involved in the control of sexual behaviours (Garelick and Swann, 2014). Essentially, sex hormones affect all levels of neuronal and neural differentiation, depending on the cell type expressing various steroid receptors, depending on the tissue, and on the sex.

The songbird's brain undergoes enormous changes induced by the effect of sex hormones. The neuronal morphology including synaptic density and dendritic arborization, the neuron number and the vascularization of the song nuclei are regulated

by sex hormones (reviewed in Chen et al., 2013; Gahr, 2004). The changes in the male canary song system that are produced seasonally in parallel to changes in the singing behaviour are correlated to circulating testosterone levels (Nottebohm et al., 1987; Voigt and Leitner, 2008; Tramontin and Brenowitz, 2000).

AR expressing cells were found present in regions of the brain that are conserved among vertebrate species or linked to taxa-specific sexual behaviors (Kim et al. 1978). In songbirds in particular, androgen receptor expressing cells are present in areas of the brain related to song learning and song production such as HVC, RA, lMAN, mMAN, the auditory nucleus NIf, area X, the medial preoptic area and the medial amygdala. (Hutchison and Steimer 1984; Alward et al. 2013; Cordes et al. 2015; Balthazart et al. 1992; Bernard et al. 1999; Gahr et al. 1998, 2008; Metzdorf et al. 1999; Fusani et al. 2000; Voigt and Gahr 2011; Fraley et al. 2010; Quispe et al. 2016; Balthazart et al. 1992; Metzdorf et al. 1999; Fusani et al. 2000).

Androgen receptor structure

The androgen receptor (AR) is responsible for the genomic mechanisms activated after its interaction with testosterone or 5α-dihydrotestosterone (DHT), one of testosterone metabolites to which it has a high affinity (Grino et al., 1990). The human AR gene is formed by eight protein-coding exons (Figure 11). Exon 1 encodes the amino terminal domain (NTD) (Ferro et al., 2002; Ding et al., 2004, 2005), exon 2 and 3 the DNA binding domain (DBD) (Shaffer et al. 2004), exon 4 the hinge region (HR) and exons 5 to 8 the ligand binding domain (LBD) (Matias et al., 2000). It has two transactivation regions, one ligand independent, AF-1, located in the NTD（Callewaert et al., 2006), one ligand dependent, AF-2, located in the LBD（Slagsvold et al., 2000). Furthermore, AF-1 contains two transcription activation units Tau-1 and Tau-5. Interestingly Tau-5 retains the activation potential in the absence of the LBD while Tau-1 depends on it to be activated（Jenster et al., 1995). The AR also contains a nuclear localization signal (NLS) that imports the receptor into the nucleus and a nuclear export signal (NES) responsible for moving the receptor back to the cytoplasm in case of ligand withdrawal (Tan et al., 2015). The sequences of DBD, HR, and LBD are highly conserved among vertebrates, whereas the NTD is less conserved. The AR gene also has a variable number of untranslated regions (5′ UTRs) and as such might have additional promoters next to the

promoter adjacent to the transcription start site (Frankl-Vilches et al., 2017). In the avian AR gene putative promoters were identified in front of exons 1 and 2.

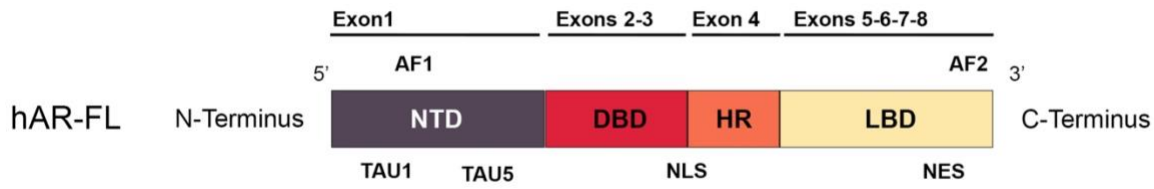


Figure 11. Androgen receptor (AR) structure. The AR has an amino terminal domain (NTD) formed by the exon 1 and containing the transcription activating units Tau1 and Tau5 within the transactivation region AF-1. A DNA binding site (DBD) formed by exons 2 and 3, a hinge region (HR) formed by exon 4 and a ligand binding domain (LBD) formed by exons 5 to 8 and containing a transactivation region AF-2. A nuclear localization signal (NLS) and a nuclear export signal (NES) can be found between the DBD and the H and in the LBD respectively.

Mechanism of action of the androgen receptor

The AR has two mechanisms of action: a canonical AR signalling and a non-canonical AR signalling. In the absence of ligand the AR is located in the cytoplasm, when it binds to testosterone or DHT a conformational change is produced and the AR/Androgen complex translocate to the cell nucleolus. Once in the nucleolus the AR dimerizes and binds to androgen response elements (AREs) present in AR target genes to modulate their expression (Denayer et al., 2010). Coregulators can bind to the activated AR to enhance or repress its effect (Bevanet al., 1999). This process constitutes the canonical signalling. Unlike the canonical signalling, the non-canonical signalling has a DNA binding independent mechanism of action, activating other second messenger signalling pathways such as ERK, AKT and MAPK (Figure 12) (Estrada et al., 2003; Kousteni et al., 2001; Gill et al., 2004).

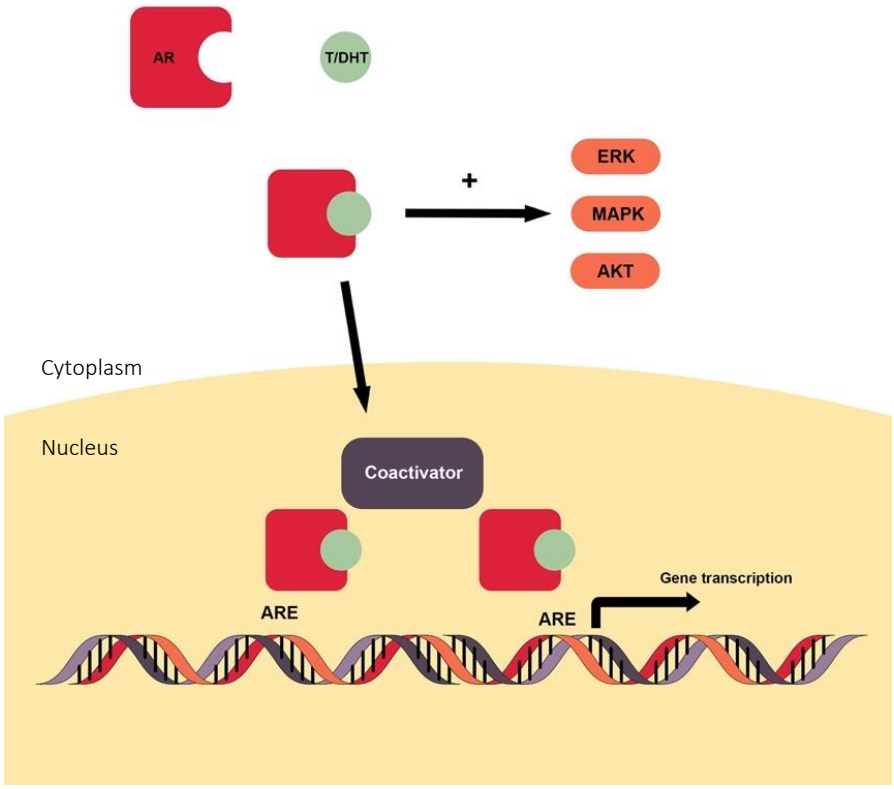


Figure 12. AR acts as a nuclear transcription factor. The AR is initially located in the cytoplasm and after binding to its ligands, testosterone (T) or dihydrotestosterone (DHT), the AR/ligand complex is transported into the nucleus. Once there it interacts with Androgen Response Elements (ARE) of a gene and modulates its expression, this constituting the canonical AR signalling pathway. Non-canonical AR signalling leads to the activation of signalling pathways such as the ERK-, AKT- and MAPK-pathways.

Androgen receptor splice variants

Alternative splicing is a mechanism that occurs to pre-mRNA sequences of a gene leading to an increase of its functional diversity. Androgen receptor splice variants were first identified in humans in 2001 in prostatic cancer cell lines (Gregory et al., 2001, Heile et al., 2011, Lu et al., 2013, Ahrens-Fath et al. 2005). Later, in 2005, a natural occurring variant, the AR45 was identified in human placenta and found to be expressed also in

heart, muscle, uterus, prostate, lung, and breast (Figure 13). The AR45 lacks the exon one usually responsible for the transactivation effect, but expresses part of the sequence located between exons 1 and 2 called exon 1b followed by exons 2 to 8 (Figure 13). AR45 acts as a negative regulator of the AR signalling by inhibiting its activity when binding to androgens in the nucleus and competing with the full-length functional AR. However, AR45 also possesses transcriptional activity under overexpression of the AR co-activators b-catenin or TIF-2 (Ahrens-Fath et al., 2005). Additional natural occurring variants ARV1, 3, 7 and 9 were identified recently they lack the hinge region and ligand binding domain and contain the cryptic exons 1, 3, 4 or 5 (Figure 13, Hu et al., 2014). Studies done on the androgen insensitivity syndrome and prostate cancer have identified several more pathological AR splice variants formed by different combination of the AR exons resulting in a loss or gain of function (reviewed by Wadosky and Koochekpour, 2017). Naturally occurring AR splice variants and their physiological function have not yet been analysed in avian species.

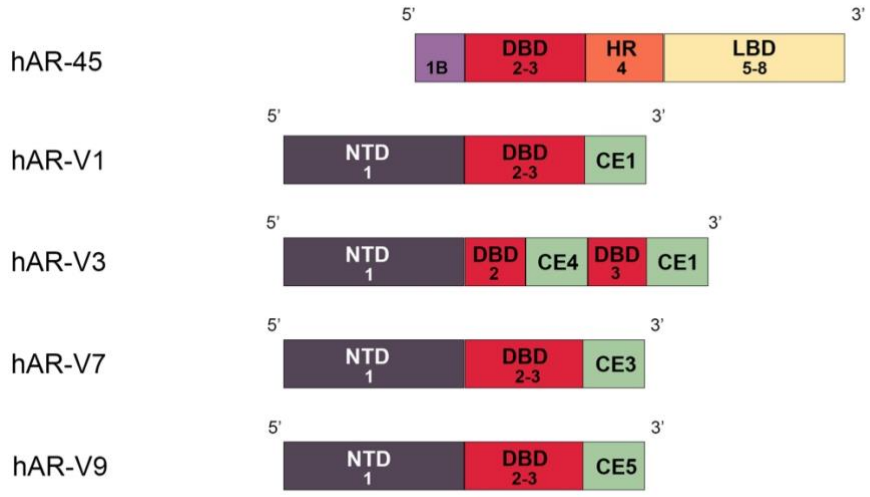


Figure 13. Human full-length androgen receptor and splice variants. HAR-FL represents the full length human androgen receptor with an amino terminal domain (NTD) formed by the exon 1 and containing a Tau1 and Tau5 transactivation sequences within the transactivation region AF-1. A DNA binding site (DBD) formed by exons 2 and 3, a hinge region (HR) formed by exon 4 and a ligand binding domain (LBD) formed by exons 5 to 8 and containing a transactivation region AF-2. A nuclear localization signal (NLS) and a nuclear export signal (NES) can be found between the DBD and the H

and in the LBD respectively. The LBD, HR and DBD are highly conserved even between mammals and birds while the NTD is less conserved in vertebrates. hAR45 constitutes the most highly expressed hAR natural occurring splice variant that it differs from the full-length AR only in the exon 1 since hAR45 contains the exon 1B found between exons 1 and 2. hARV 1, 3, 7 and 9 are less frequently expressed splicing variants that lac the hinge region and ligand binding domain and contain the cryptic exons 1, 3, 4 or 5.

Dominant-negative AR as a candidate gene for the generation of a transgenic knock-down zebra finch model

Sex hormones affect the development, organization and plasticity of the song control system and modulates the singing behaviour. In order to elucidate the androgenic effects, I involved the method developed in chapter 1 of my thesis to produce transgenic songbirds opening the possibility for the development of an AR transgenic knock-down zebra finch model. This can be achieved by the overexpression of a dominant negative variant of the androgen receptor in the genome of a transgenic zebra finch. A dominant negative variant would compete with the native AR in its ability to bind to androgens but will not produce an activation cascade, leading to a reduction of the effect of androgens in the body. Therefore, for chapter 3 of my thesis I've we studied naturally occurring AR splice variances of the zebra finch and tested their activity *in vitro* in a luminescent promoter assay. Furthermore, I've tested the zebra finch homolog of a synthetic dominant negative variant of the human AR variant whether it could be used in the generation of a transgenic AR knock-down zebra finch model.

Thesis outline

Chapter I

Over the last 15 years, three transgenic zebra finch models were developed using lentiviral vector injection into the blastoderm of songbirds. To achieve genetic manipulations of songbirds with higher efficiency, in the first chapter of this thesis I have

developed a new method by the culture and genetic modification of primordial germ cells (PGCs) (Gessara et al., 2021). Novel techniques were developed to expand zfPGCs *in vitro* for a period of time sufficient for their genetic modification with lentiviral vectors. Genetically modified PGCs were injected into early zebra finch embryos and transgenic offspring was produced in the second generation.

Chapter II

In order to apply precise genome editing tools like CRISPR/Cas9, it seemed necessary to culture the PGCs for an extended period of time. Because that allows for the application of gene editing techniques followed by antibiotic selection and clonal expansion of the targeted cells. Long-term PGC cultures have only been established for chicken PGCs so far. zfPGCs can be cultured for a relative short period of time that does not allow for the clonal selection. Since the growth curve of zfPGCs in culture reached the stationary phase between 10 and 15 days in vitro (DIV), a comparative transcriptome study of zfPGCs cultured for 1 and 10 DIV was performed in this chapter. The goal was to determine the signalling pathways affected in zfPGCs during culturing and to identify growth factors potentially necessary to maintain long-term zfPGC self-renewal. Furthermore, candidate growth factors were tested on their effect on the proliferation rate of the zfPGCs *in vitro*.

Chapter III

Genetic manipulation of cultured zfPGCs can be used for the generation of transgenic zebra finches with an over-expression of a transgene (Gessara et al., 2021). In chapter III I have studied the expression and activity of androgen receptor (AR) variants to identify a candidate transgene that can be used for the generation of a transgenic zebra finch expressing a dominant-negative AR-variant. In birds, the full-length androgen receptor constitutes a key regulator of many developmental and behavioural processes including the development of the song control system in songbirds (Gahr, 2004; Chen et al., 2013). I have investigated here the expression pattern of two novel zebra finch AR splice variants zfAR-V2-4 and zfAR-V2-8 by RT-PCR and assessed their genomic activity in an in vitro AR reporter assay. The results suggested that the naturally occurring zfAR-V2-8 constitute a viable tool for the generation of a transgenic zebra finch with an AR knock-

down. The availability of an inhibitory AR variant opens the possibility to develop a transgenic dominant-negative knock-down model in the zebra finch to study the role of androgens in the development, organization and plasticity of the song control system and the modulation of the singing behaviour.

Chapter I

Highly efficient genome modification of cultured primordial germ cells with lentiviral vectors to generate transgenic songbirds

Ivana Gessara

Falk Dittrich

Moritz Hertel

Staffan Hildebrand

Alexander Pfeifer

Carolina Frankl-Vilches

Mike McGrew

Manfred Gahr

Published in Stem Cell Reports (2021)

Highly efficient genome modification of cultured primordial germ cells with lentiviral vectors to generate transgenic songbirds

Authors

Ivana Gessara[1]*, Falk Dittrich[1], Moritz Hertel[1], Staffan Hildebrand[2], Alexander Pfeifer[2], Carolina Frankl-Vilches[1], Mike McGrew[3], Manfred Gahr [1]

1. Max Planck Institute for Ornithology, Eberhard Gwinner Strasse, 82319 Seewiesen, Germany

2. Institute of Pharmacology and Toxicology University of Bonn, Venusberg-Campus 1, 53127 Bonn

3. The Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian EH25 9RG, United Kingdom

Abstract

The ability to genetically manipulate organisms has led to significant insights in functional genomics in many species. In birds, manipulation of the genome is hindered by the inaccessibility of the one-cell embryo. During embryonic development, avian primordial germ cells (PGCs) migrate through the blood stream and reach the gonadal *anlage*; where they develop into mature germ cells. Here, we explored the use of PGCs to produce transgenic offspring in the zebra finch, which is a major animal model for sexual brain differentiation, vocal learning and vocal communication. Zebra finch PGCs (zfPGCs) obtained from embryonic blood significantly proliferated when cultured in an optimized culture medium and conserved the expression of germ and stem cell markers. Transduction of cultured zfPGCs with lentiviral vectors was highly efficient leading to strong expression of the enhanced green fluorescent protein (eGFP). Transduced zfPGCs were injected into the host embryo and transgenic songbirds were successfully generated.

Introduction

Songbirds are major animal models to study the genetic and neural basis of vocal learning and communication (Mooney, 2020; Mello, 2014; Prather et al., 2017) as well as sex hormone dependent brain development (Gahr, 2007; Balthazart et al., 2010; McCarthy and Arnold, 2011) and adult neurogenesis (Goldman and Nottebohm, 1983; Paton and Nottebohm, 1984). For the zebra finch, transgenic models have been successfully developed (Agate et al., 2009; Abe et al., 2015; Liu et al., 2015). Agate et al. (2009) were the first to inject lentiviral vectors for GFP into the blastodisc of freshly laid zebra finch eggs in order to target the primordial germ cells (PGCs), the precursors of spermatocytes and oocytes. However, due to inefficiency of the method, only two other transgenic models have been generated over the past 10 years using this method (Abe et al., 2015; Liu et al., 2015).

During avian development, PGCs are located in the central area of the blastodisc until Eyal-Giladi and Kochav (EGK) stage X (Eyal-Giladi et al., 1976; Kochav et al., 1980); from here, PGCs translocate to the germinal crescent (Ginsburg et al., 1986) and finally migrate through the developing vascular system to reach the gonadal *anlage* (Fujimoto et

al., 1976; Nakamura et al., 1988; Jung et al., 2019). Avian PGCs are easily isolated from the circulatory system by blood aspiration between Hamburger and Hamilton (HH) stages 14 and 17 (Hamburger and Hamilton, 1951). In the domestic chicken, embryonic blood-derived PGCs can be propagated *in vitro* for several months, genetically modified and re-injected into early embryo surrogate hosts. Their ability to migrate to the gonadal *anlage* was unaffected by this treatment and the resulting hosts exhibited a high germline transmission rate (Van De Lavoir et al., 2006; Macdonald et al., 2010). For the zebra finch, it has been shown that primordial germ cells of the embryonic gonads (zfgPGCs) can be cultured for up to 30 days on a feeder cell layer in the presence of relative high concentrations of undefined fetal bovine serum, genetically modified with transposons and retain the ability to colonize embryonic host gonads (Jung et al., 2019). However, to the best of our knowledge the generation of transgenic songbirds using this approach has not yet been reported.

Here, we explored the use of blood-borne zfPGCs to generate transgenic zebra finches using a feeder-free culture medium that contained fetal bovine serum at a minimal concentration to avoid precocious differentiation and using lentiviral transduction for genetic modification. Cultured zfPGCs were efficiently transduced with lentiviral vectors expressing enhanced green fluorescent protein (eGFP). After injection underneath the early zebra finch blastodisc of a host embryo (EGK stage X), genetically modified zfPGCs migrated to and colonized the gonadal *anlage* following incubation. These surrogate host embryos were successfully hatched and raised by foster parents and, when mated, efficiently transmitted the transgene to their offspring (Fig. S1).

Results

Expansion of embryonic blood-derived zfPGCs in culture

Blood samples (1 - 3 µl) extracted from zebra finch embryos at Murray stages 13-15 (Murray et al., 2013) contained a mixed population of erythrocytes and about 140 PGCs, with large individual differences in the number of PGCs observed between samples (40 – 500). Like in chicken (Macdonald et al., 2010), blood-derived zfPGCs were round with a

diameter of 16-20 μm, showed tiny membrane protrusions visible under phase contrast, a prominent nucleus and small cytoplasmic vesicles that enclosed polysaccharides which could be stained by the periodic acid-Schiff (PAS) reaction (Fig. 1 A-C).

To culture zfPGCs, we used a feeder layer free culture and a defined culture medium (FAIcs) with low calcium concentration that we had developed for chicken PGCs (chPGCs) obtained from embryonic blood (Whyte et al., 2015), but replaced the chicken serum with a low concentration of fetal bovine serum. In this culture medium, we found that a low concentration of calcium (0.3 mM) supported the proliferation of zfPGCs in loose, non-adherent clumps (Fig.1 D, Fig.S2 A). Higher concentrations of calcium (medium with 100 % KnockOut^{TM} DMEM containing 1.8 mM calcium) promoted stronger cell-cell interactions, leading to the formation of dense zfPGC clumps with indistinguishable cell boundaries. We also tested a culture medium previously developed for zfgPGCs (Jung et al., 2019) that contained higher serum concentrations. Using this medium condition, zfPGCs extracted from embryonic blood attached to the bottom of the culture plate (Fig. S2 B), showed a lower growth rate and exhibited signs of cell death.

To establish improved culture conditions for zfPGCs, we adapted the FAIcs by altering the ligand of the TGF-beta signaling pathway. Both growth factors Activin A and Bone morphogenetic protein (BMP) 4, have been shown to be sufficient for propagating chicken PGCs (chPGCs) *in vitro* (Whyte et al., 2015). However, we found a trend towards higher numbers of cell divisions in the presence of BMP4 for zfPGCs (average and SEM: 4.33 ± 0.88 for Activin and 6.97 ± 1.87 for BMP4; n=3; Fig. 2E), and replaced Activin with BMP4 for our germline transmission experiments. Moreover, zfPGC cultured in the presence of BMP4 had after 15 days in culture a significantly higher percentage of cells than the control without BMP4 (average percentage 206%; SEM 36.3; p= 0.05; n=5).

When our optimized zfPGC culture medium was used (Table S1), proliferation of zfPGCs was detectable after 2 days *in vitro* (DIV), and loose cell clusters were observed forming after 5 DIV (Fig. 1D). During the first two weeks of culture, zfPGCs divided 2 to 9 times, reaching a maximum of 8000 cells. On average, we obtained 1331 zfPGCs at 10 DIV and 1045 zfPGCs at 15 DIV per blood sample (n = 32; Fig.1 F, G). Between 15 and 20 DIV, zfPGCs started to undergo cell death.

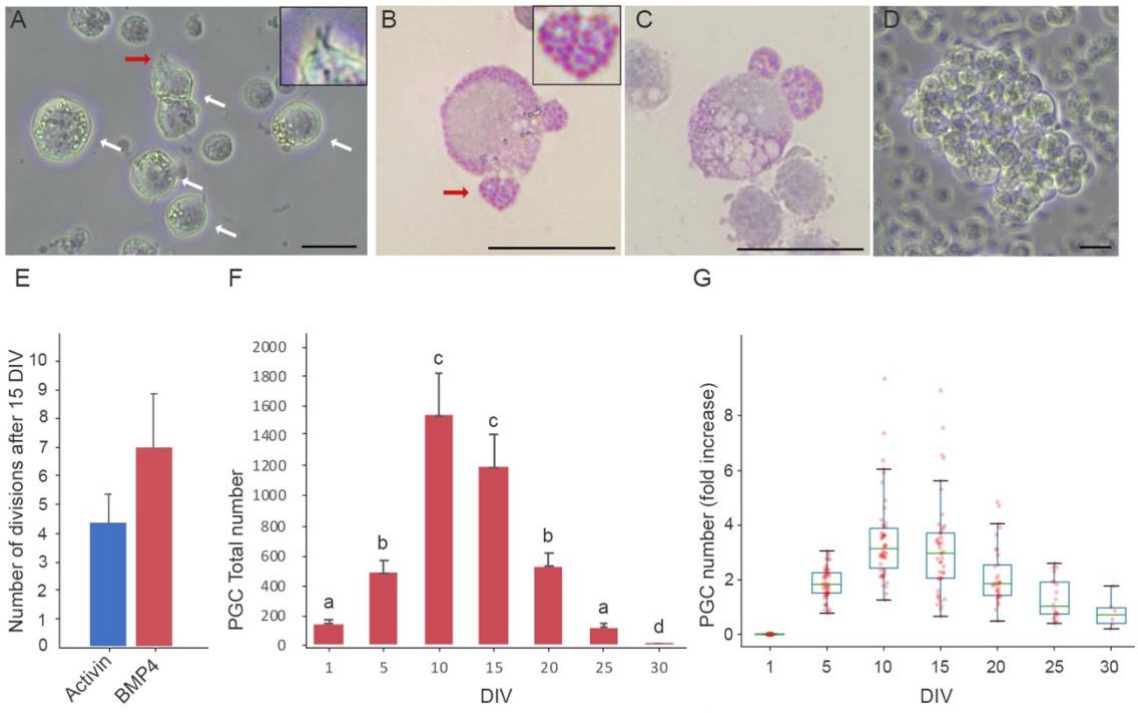


Figure 1. Characterization and expansion of cultured zfPGCs

(A) Phase contrast image of an early blood cell culture with zfPGCs (white arrows) that contained polysaccharide vesicles and showed typical cellular protrusions (red arrow). The insert in A shows cellular protrusions at higher magnification. Note that PGCs were larger in size than other blood cells. **(B, C)** Periodic acid-Schiff (PAS) stains of blood smear samples taken from a zebra finch **(B)** and chicken **(C)** embryo at embryonic stage 14, staining polysaccharide vesicles of PGCs in magenta. A cytoplasmic fragment of the zfPGC (red arrow) is shown in the insert at a higher magnification. **(D)** ZfPGCs growing in loose clumps at 7 DIV conserve their native morphology. **(E)** Blood samples from single embryos were split in two and cultured one half in the presence of Activin and the other with bone morphogenetic protein 4 (BMP4). Cell numbers were counted at 1 and 15 DIV and numbers of cell divisions at 15 DIV are shown (mean ± SEM). **(F)** Growth curve of zfPGCs cultured in medium containing BMP4. Different letters indicate statistically significant differences between cell numbers on different days (ANOVA; n=32, p=0.05). ZfPGCs proliferated significantly until 15 DIV. At 20 DIV a significant decrease of the cell number due to death was detectable. **(G)** Box plot showing fold

increase in zfPGCs during a period of 30 days in culture as compared to the number of zfPGCs at 1 DIV. Scale bars in A - D represent 20 μm. See also Figure S2

Expression of cell type-specific markers by cultured zfPGCs

To test the effect of *in vitro* propagation on PGC-specific gene expression we assessed germ and stem cell marker expression by zfPGCs at 10 DIV by performing immunofluorescent stainings and RT-PCR analyses. Stage-specific embryonic antigen 1 (SSEA-1), a carbohydrate epitope associated with cell adhesion and migration, is normally expressed by avian PGCs during migration to the gonadal *anlage* in turkeys (D´Costa and Petitte, 1999), as well as in zfgPGCs (Jung et al., 2019) and also by chPGCs (Macdonald et al., 2010) *in vitro*. Furthermore, epithelial membrane antigen 1 (EMA-1), a cell surface glycoprotein, is expressed by chPGCs at early embryonic stages and *in vitro* (Urven et al., 1988; Raucci et al., 2015). Correspondingly, migrating zfPGCs that were freshly isolated from the embryonic blood stream at Murray stages 13-15 as well as zfPGCs at 10 DIV were found to be immunopositive for both SSEA-1 (Fig. 2 A, D) and EMA-1 (Fig. 2 E, H, I). In both antibody stainings, the cell surface of zfPGCs was strongly labeled (Fig. 2, D); in contrast, erythrocytes remained unstained (Fig. 2 A, B, E, F).

In both chPGCs and zfPGCs, the germline-specific genes DAZL (Deleted in Azoospermia Like) and DDX4 (DEAD-Box Helicase 4) as well as the pluripotency markers POU5F1 (POU Class 5 Homeobox 1) and NANOG (Nanog Homeobox) are expressed (Van De Lavoir et al., 2006; Macdonald et al., 2010; Jung et al., 2019). Our RT-PCR analyses revealed that cultured zfPGCs harvested after 10 DIV expressed these stem cell markers (Fig. 2 J). In control samples, zebra finch blastodiscs from freshly laid eggs, which contain PGCs, showed a similar expression pattern. In contrast, cultured fibroblasts from zebra finch embryos did not express these markers (Fig. 2 J). Furthermore we identified the PGC marker DAZL to be highly expressed in the transcriptome of both zfPGCs freshly extracted from blood and cultured for 10 DIV.

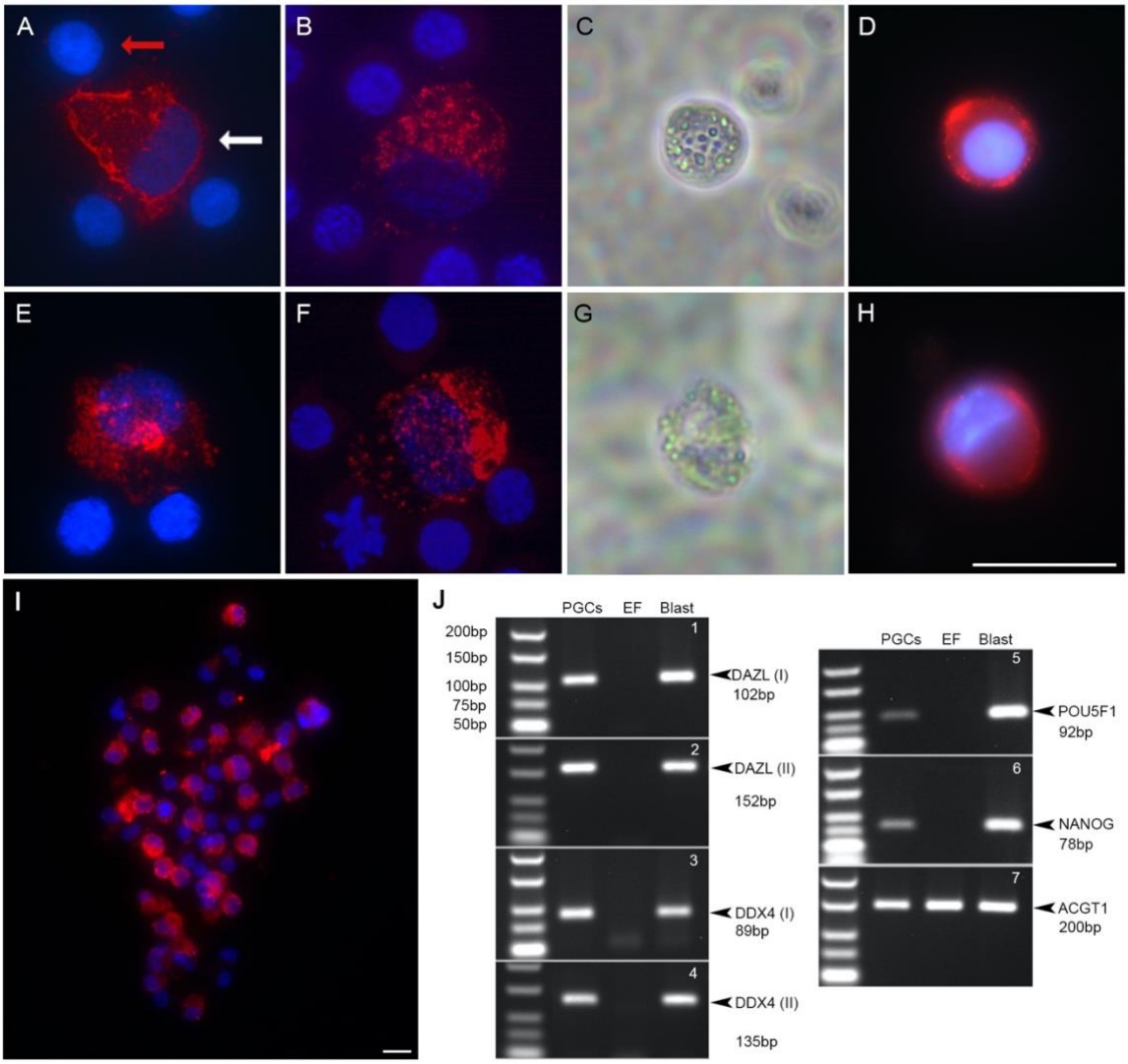


Figure 2. Conserved expression of germ and stem cell-specific markers in cultured zfPGCs

Microphotographs of PGCs were taken after immunofluorescent stainings for **(A - D)** stage-specific embryonic antigen 1 (SSEA1) and **(E - I)** epithelial membrane antigen 1 (EMA1). Stainings were performed with blood smears obtained from zebra finch (zf) (**A, E**) and chicken (ch) **(B, F)** embryos as well as with zfPGCs cultured for 10 DIV **(C, D, G, H, I)**. Fluorescent images (immunostainings in red and nuclear stains with 4′,6-diamidin-2-phenylindol (DAPI) in blue) and phase contrast (C, G) images are shown. In (A) arrows point to a zfPGC (white arrow) and an erythrocyte (red arrow). Note that in blood smear samples PGCs were surrounded by immunonegative blood cells. Scale bar represents 20µm. **(J)** Photographs of agarose gels after electrophoretic separation and ethidium bromide staining of RT-PCR products. Cultured zfPGCs (PGCs; 10 DIV)

continued to express germ cell markers such as DAZL (1, 2) and DDX4 (3, 4), and, interestingly, also stem cell markers such as POU5F1 (5) and NANOG (6). In contrast, cultured fibroblasts of zebra finch embryos did not express any of these marker genes. All cDNA samples from zebra finch blastodisc cells (Blast) that include early zfPGCs turned out to be PCR-positive for the same marker genes. Detection of y-actin (ACTG1) (7) expression was used to control all samples for equal cDNA quality. Sizes of molecular weight markers (left slots) and expected amplicons are indicated. For both DAZL and DDX4 two RT-PCRs were carried out using primer pairs that were reported by Mak et al. 2015 (I) and newly designed for this study to be intron-spanning (II) (see in material and methods).

Highly efficient transduction of cultured zfPGCs with lentiviral vectors

Electroporation (BTX, Gemini System) and lipofection resulted in extensive cell mortality of cultured zfPGCs and low transfection efficiency (data not shown). To determine whether cultured zfPGCs could be genetically modified, we transduced cultured zfPGCs at 7 DIV with a lentiviral vector containing eGFP gene under control of the human phosphoglycerate kinase (hPGK) promoter (Fig. 3 A) when they were growing in cell clumps. Two days post-transduction, the majority of the cells exhibited strong eGFP expression (Fig. 3 B-D) while continue to express the zfPGC marker EMA1(Fig. 3 E). In contrast to zfPGCs cultured from embryonic blood, zfPGCs cultured from embryonic gonads (zfgPGCs) (Jung et al., 2019) showed a relatively low lentiviral transduction efficiency (Fig. S2 G, H). We next compared the eGFP expression in cultured zfPGCs after transduction with lentiviral vectors that contained different constitutive promoters including the hUBC promoter, the human elongation factor 1 alpha (hEF1α) promoter, the cytomegalovirus (CMV) promoter and the CMV enhancer fused to the chicken beta-actin (CAG) promoter. For all these lentiviral constructs the level of the eGFP expression after transduction turned out to be similar (Fig. S3).

Since viral titer is a crucial factor for a successful lentiviral vector transduction, we also assayed reporter gene expression in cultured zfPGCs when using increasingly diluted viral particle suspensions. Using the lentiviral vector for rrl-hPGK-eGFP, we found the highest eGFP expression at a final concentration of 2x10exp8 TU/ml and greater (Fig. S4). For gene editing studies in which Cas9, CRISPR gRNAs and a selection marker are

used at the same time, it would be necessary to simoultaneously transduce zfPGCs with two different lentiviral vectors. To test this we treated the same cells with lentiviral vectors for both UBC-eGFP and CMV-Tomato at a titer of 2x10exp8 TU/ml each. Two days after transduction, most zfPGCs expressed both reporter genes at comparable strength (Fig. S5). Together, these findings demonstrated that lentiviral vectors provide an efficient and promising way to introduce transgenes into cultured zfPGCs.

Upregulation of LDLR gene family members in cultured zfPGCs

The viral vectors used to transduce zfPGCs in this study were vesicular stomatitis virus glycoprotein G (VSV-G) pseudotyped lentivirus for which the low-density lipoprotein receptor family (LDLRF) constitutes the main cell surface receptors (Finkelshtein et al., 2013, Nikolic et al., 2018). Transcriptome analysis of zfPGCs that were freshly extracted from embryonic blood and cultured for 10 DIV revealed the expression of the LDLR gene family members LR11, LRP1B and LRP3 to be significantly up-regulated after 10 DIV. The expression of other gene family members such as LRP2, LRP6, LRP8 and LRP1 remained unchanged and was found to be down-regulated for VLDLR and LRP4 (Table 1). Absence of the prototypic LDLR gene in the zfPGC transcriptome was in line with the role of LDLR in steroidogenesis by somatic cells of chicken ovarian follicles (Hummel et al., 2003).

Table 1. Expression profile of Low density lipoprotein receptor family in zfPGCs cultured for 10 days

Gene Symbol	Gene Name	Zebra finch Ensembl gene ID	Fold Change	Adjusted P. value
LR11 (SORL1)	Low density lipoprotein receptor relative with 11 ligand binding repeats	ENSTGUG00000000462	2.816	0.05
LRP1B	Low density lipoprotein receptor related protein 1b	ENSTGUG00000011849	2.602	0,043
LRP3	Low density lipoprotein receptor related protein 3	ENSTGUG00000009422	1.8494	0.008
LRP2 (Megalin)	Low density lipoprotein receptor related protein 2	ENSTGUG00000007663	1.172	0.175
LRP6	Low density lipoprotein receptor related protein 6	ENSTGUG00000012717	0.412	0.698
LRP8 (APOER2)	Low density lipoprotein receptor related protein 8	ENSTGUG00000009223	-0.008	0.994
LRP1	Low density lipoprotein receptor related protein 1	ENSTGUG00000015747	-0.560	0.620
VLDLR	Very low density lipoprotein receptor	ENSTGUG00000005386	-1.558	0,017
LRP4 (MEGF7)	Low density lipoprotein receptor related protein 4	ENSTGUG00000010533	-2.688	0,004

Cultured zfPGCs colonized the gonadal anlage of host embryos after re-injection subgerminally

Cultured zfPGCs clumps expressing eGFP under control of the hPGK promoter were pooled from 5 to 10 cultures and dissociated to single cells after digestion with papain. Around 500 cells were injected under the blastodisc of freshly laid zebra finch eggs. The eggs were sealed and incubated for 48 hours before being transferred into the nests of

foster parents for further incubation. From 22 injected eggs 10 founder birds (45,4 %) were hatched and raised, six females and four males. Using histological sections of founder gonads, we found that cultured and re-injected zfPGCs migrated to the gonadal anlage of the host embryo and differentiated into germinal cells (Fig. 3 G, H). Furthermore, integration of the hPGK-eGFP construct in genomic DNA extracted from founder gonads was verified by PCR in both ovaries and testes (Fig. 3 I). All founders that were examined contained hPGK-eGFP-positive zfPGCs in their gonads.

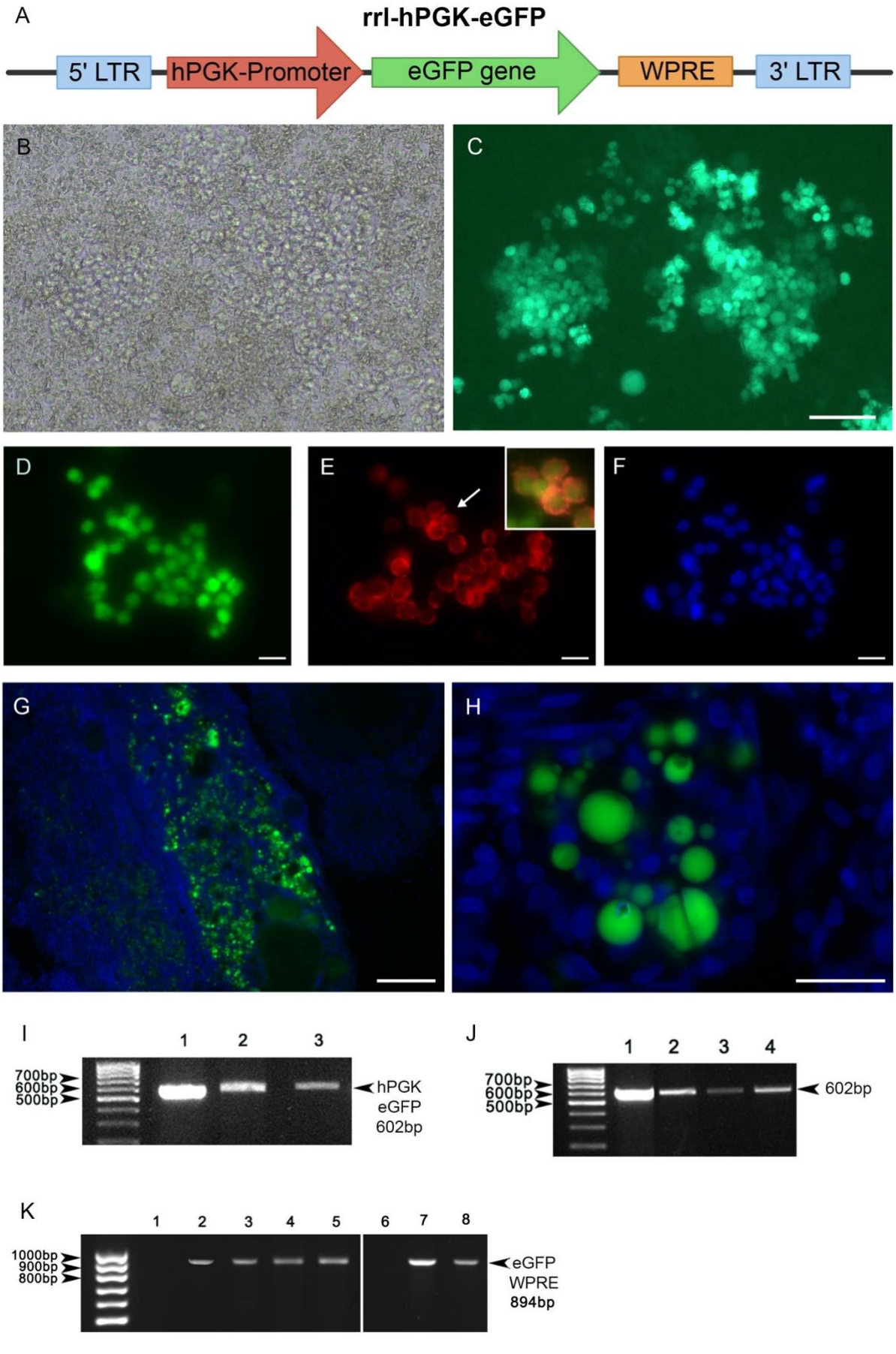

A
rrl-hPGK-eGFP
5' LTR
hPGK-Promoter
eGFP gene
WPRE
3' LTR
B
C
D
E
F
G
H
I
1
2
3
700bp
600bp
500bp
hPGK
eGFP
602bp
J
1
2
3
4
700bp
600bp
500bp
602bp
K
1
2
3
4
5
6
7
8
1000bp
900bp
800bp
eGFP
WPRE
894bp

Figure 3. Cultured zfPGCs were efficiently transduced with a lentiviral vector for eGFP and colonized host embryo gonads

(A) Lentiviral vector rrl-hPGK-eGFP that was used to transduce zfPGCs *in vitro* and generate founder birds. **(B-F)** Microphotographs of cultured zfPGCs that were transduced with rrl-hPGK-eGFP for 48 hours. Phase contrast **(B)** and fluorescent images **(C, F)** are shown for zfPGCs expressing eGFP **(C, D)** and the zfPGC marker EMA1 **(E)** in E, a group of EMA1 positive zfPGC (white arrow) is shown in the insert merged with eGFP from D**.** zfPGCs nucleus were counterstained with DAPI **(F).** In D-F arrows point to the same zfPGC**.** Note that almost all zfPGCs were expressing the reporter gene. In **(F, G)** microphotographs of cryosections produced from an adult founder ovary that showed eGFP expression and blue nuclear counterstain with DAPI are presented at low (**F**) and high (**G**) magnification. Note that in the ovary eGFP-positive germ cells of different sizes display various stages of differentiation. Scale bar represents B,C, G, H 100 μm, D-F 20 μm. Genomic DNA (gDNA) was isolated from (I) founder gonads and (J) blood samples of transgenic F1 birds to detect a hPGK-eGFP sequence (602 bp) of the lentiviral construct by PCR. After electrophoretic separation images of agarose gels are shown that contained a molecular weight marker (left lanes) and PCR products that were obtained using the lentiviral transgene plasmid (I, 1; J, 1), gDNA isolated from a founder ovary (I, 2) and testis (I, 3), and blood gDNA of three F1-birds, two males (J, 2 and 3) and one female (J, 4). (K) RNA was isolated from brain (K, 2 and 3) and liver (K, 4 and 5) and gDNA was isolated from muscle (K, 7 and 8) of eGFP-transgenic birds. K 1 and 6 are negative RNA and gDNA controls respectively. See also Figures S3-S5.

Songbird transgenesis

Founders were crossed with wild type birds, and genomic DNA from blood samples of the offspring (F1 birds) were analyzed by PCR for the presence of the hPGK-eGFP construct (Fig. 3 J). F1 birds that were hPGK-eGFP-positive by PCR were sacrificed to produce histological sections and confirm eGFP-protein expression by anti-GFP immunofluorescence staining. All F1 birds that were found to be hPGK-eGFP-positive by PCR turned out to be GFP-immunopositive as well. Adittionally RNA samples were obtained from liver, brain and muscle of positive F1 birds and RT-PCR was performed to identify the eGFP RNA. GFP was found to be present in RNA samples from brain and

liver but not from muscle. However, the hPGK-eGFP gene was found present when performing PCR from F1 muscle Genomic DNA, suggesting that the GFP gene is ubiquitously integrated but not ubiquitously expressed (Figure 3 K). In summary, all 10 founder birds produced transgenic offspring showing a germline transmission rate between 4 and 22 % (Table 2). In the transgenic birds, we observed eGFP expression in kidney, heart, lung, liver, ovary and brain (Fig. 4), including forebrain song control nuclei like HVC (Fig. 5 A-C) and the robust nucleus of the arcopallium (RA; Fig. 5 D-F). The percentage of eGFP expressing cells varies from 1 to 70% depending on the organ and the individual (Table S2). All transgenic birds generated were phenotypically normal and did not present any pathologies (Fig. 5 G).

Table 2. Frequency of germline transmission detected by PCR and anti-GFP immunostaining

Founders		F1 birds			
ID	Sex	Number	hPGK-eGFP-positive by PCR	GFP-immunopositive	Germline transmission (%)
1	Male	18	4	4	22
2	Male	8	1	1	12
3	Female	9	1	1	11
4	Female	11	1	1	9
5	Male	23	2	2	8
6	Male	13	1	1	7
7	Female	17	1	1	5
8	Female	19	1	1	5
9	Female	20	1	1	5

10	Female	23	1	1	4

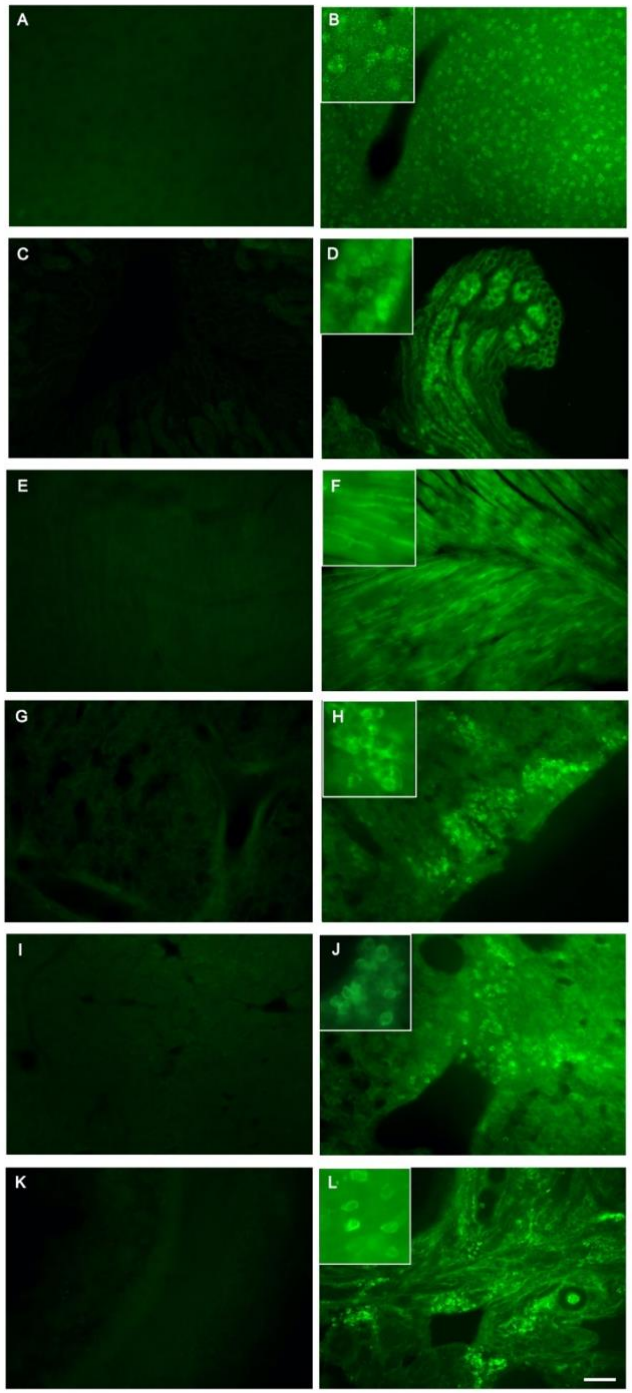


Figure 4. EGFP expression in transgenic birds

Anti-GFP immunofluorescence stainings are shown for different organs of adult control (left panels) and eGFP-transgenic (right panels) birds: brain (**A, B**), kidney (**C,D**), heart **(E,F**), lung **(G,H)**, liver **(I,J)** and ovary **(K,L)**. Inserts show eGFP-immunopositive cells at higher magnification. The scale bar represents 100 μm.

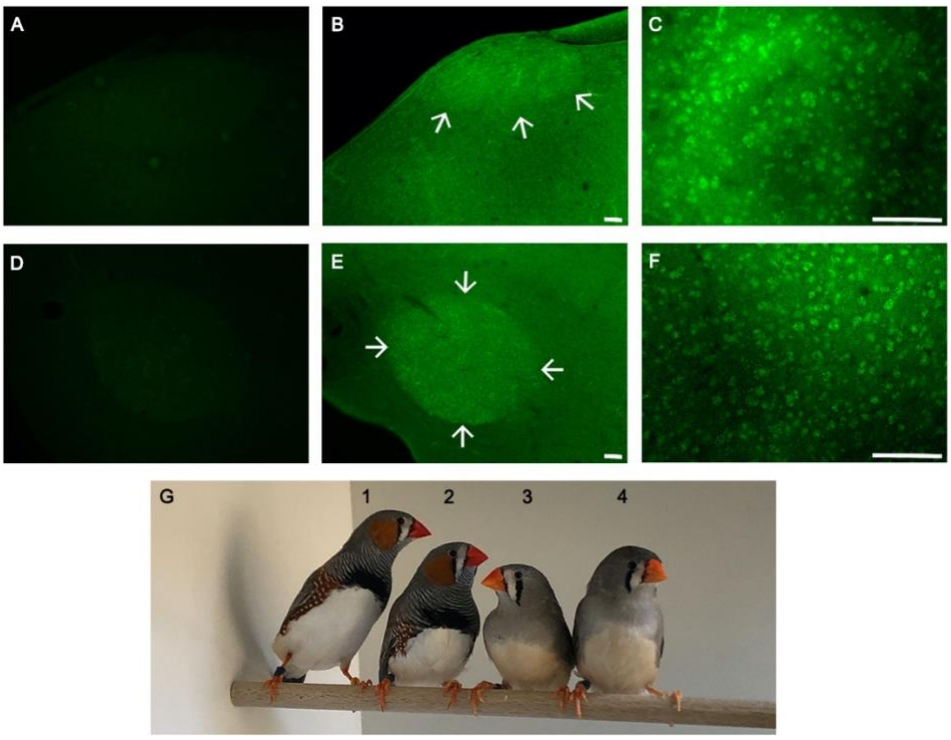


Figure 5. EGFP expression in the song control nuclei HVC and RA of transgenic birds

Anti GFP-immunofluorescence stainings are shown for adult male forebrain sections of control (**A,D**) and eGFP-transgenic birds (**B,C,E,F**) that included the song control nuclei HVC (A-**C**) and RA (D-**F**) at low (left and middle panels) and high (right panels) magnification. At low magnification location of the song control nucleus is indicated with arrows. **(G)** Photograph of adult zebra finches being transgenic (1 and 3) and wild type (2 and 4), respectively. The scale bars represent 100 μm.

Discussion

Songbirds, especially the zebra finch are prominent animal models for studying neural circuit development, vocal communication and vocal learning among other topics (Bolhuis and Gahr, 2006; Konopka and Roberts, 2016). However, the wider use of the zebra finch in mechanistic research is hampered by the lack of transgenic technology, which is available for most sandard animal model organisms (Ménoret et al., 2017). In chicken, the long-term cultures of PGCs turned out to be extraordinarily useful for facilitating the production of transgenic chicken lines (Van De Lavoir et al., 2006; Macdonald et al., 2010). The ability to culture chPGCs almost indefinitely made it possible to modify the chicken genome *in vitro* and inject large numbers of genetically modified chPGCs into host embryos substantially increasing the germline transmission rate in these birds (Sid and Schusser, 2019; Schusser et al., 2013; Schusser et al., 2016; Dimitrov et al., 2016; Taylor et al., 2017; Oishi et al., 2016; Oischi et al., 2018). Here, we adapted such an approach for the zebra finch to allow for fast and efficient generation of transgenic songbirds.

We developed a new method for the propagation of zfPGCs from early embryonic blood cultures. When grown in cell clumps after 7 DIV, zfPGCs exhibited high transduction rates using lentiviral vectors such that almost every zfPGC expressed the reporter gene. Previously, the transduction and transmission efficiency of lentiviral vectors after injection at blastodermal stages was low (Agate et al., 2009). In particular, lentiviral vectors transduced zfPGCs much less efficiently *in ovo* compared to cultured zfPGC growing in clumps. Factors like inactivation of lentiviral vectors *in ovo* and increased cellular uptake *in vitro* might be responsible for this difference. Thus, our approach constitutes a major improvement for the genomic modification of songbirds. Up-regulation of LDLR gene family members like LR11, LRP1B and LRP3 by zfPGCs may have contributed to the increased uptake of the VSV-G pseudotyped lentiviral vectors by zfPGCs growing in clumps. Interestingly, LR11 also promotes the cell adhesion process of hematopoietic stem cells (Nishii et al., 2013). We observed, that growth of zfPGCs in clumps was associated with an enhanced uptake of lentiviral vectors. Lower transduction rates we found in singled zfPGCs obtained from cultured cell clumps after digestion with papain or when prepared from embryonic gonads (Jung et al., 2019). The highly efficient transduction of cultured zfPGCs with lentiviral vectors did not compromise the ability of

the germ cells to differentiate in the host gonad and form functional gametes. All founder birds (100%) of both sexes that were generated produced transgenic offspring, which is astonishing considering the injected PGCs were derived from a mixture of male and female embryos. An additional factor that contributed to the high efficiency we achieved was the relatively high hatching success (45.4 %) of the injected host embryos. In comparison, injections of lentiviral vectors into blastoderm resulted in a hatching rate of 13.2% (Agate et al., 2009). We posit that the transfer of genetically modified zfPGCs into a host embryo by a single injection was less harmful to the embryo than the multiple vector injections required by the previous protocol used for the production of transgenic zebra finches (Agate et al., 2009). In conclusion, expansion of embryonic blood-derived zfPGCs *in vitro* followed by transduction with lentiviral vectors constitutes an extremely efficient strategy for producing transgenic zebra finches.

We selected lentiviral constructs containing the hPGK promoter to generate transgenic zebra finches because they generated consistently strong eGFP expression in cultured zfPGCs. In our transgenic hPGK-eGFP F1-birds, eGFP was not ubiquitously expressed in all zebra finch tissues but was detectable after immunostaining in most brain areas including the song control system, in the kindney, heart, lung, liver and ovary. The heterogeneous and selective expression pattern of eGFP might be the result of a cell-type specific activity or silencing of the hPGK promoter in the zebra finch. For cell-type directed transgene expression, additional promoter constructs need to be tested to match the transgene expression pattern to the specific requirements that are defined by the various experimental questions to a transgenic songbird model.

The use of cultured zfPGCs that were transduced with lentiviral vectors appears to be an efficient transgenic approach for the overexpression of a transgene, and possibly, downregulation of an endogenous gene following random transgene insertion into the genome. The successful simultaneous transduction of cultured zfPGCs with two different lentiviral vectors suggested that in future CRISPR/Cas experiments the use of high-titer lentiviral preparations for Cas-proteins under the control of a PGK-promoter and gRNAs under control of an U6-promoter will be a promising approach. During a culturing period of two weeks we expect to be sufficiently long for the successful application of targeted gene editing techniques on zfPGCs, such as gene knock-outs by non-homologous end joining (NHEJ) repair of CRISPR/Cas-induced DNA breaks (Sid and Schusser, 2018). For more complex genome editing techniques that involve a homology directed repair

(HDR), culture conditions that permit a more extended growth of zfPGCs *in vitro* and clonal selection of genetically modified zfPGCs will need to be developed. Nevertheless, for transgenic applications in neuroethological studies, the method of generating transgenic songbird models reported here represents a fast, efficient and straightforward procedure.

Experimental procedures

Ethics statement

Animal handling was carried out in accordance with the European Communities Council Directive 2010/63 EU and legislation of the state of Upper Bavaria.

Culture of embryonic blood-derived zfPGCs

Freshly laid zebra finch eggs were incubated for 60 hours at 37°C and relative humidity of 75% until Murray stage 13-15. Then a window was opened above the air chamber of the egg, and 1 - 3 µl of blood was extracted from the vasculature system using a pulled glass needle. Blood samples were cultured separately in a 96 well plate (Falcon) with 150 µl of the zfPGC culture medium in each well (Table S1). Cells were maintained in culture for up to 30 days and 80µl of the medium was replaced every other day. Gonadal PGCs (gPGCs) were obtained and cultured as described by Jung et al. (2019).

To quantify the number of cell divisions in cultures with Activin or BMP4 the number of zfPGCs was scored in a minimum of 8 fields under a 20× objective of an inverted microscope (Nikon eclipse Ts2) at 1 and 15 days of culture and the fold change between both culture periods was calculated.

Histological stainings

For periodic acid-Schiff (PAS) stain of blood smears, samples were extracted from zebra finch and chicken embryos at HH stage 14, dried on glass slides and fixed with 70%

ethanol for 15 minutes. For staining, a PAS kit (Sigma Aldrich) was used according to the manufacturer's instructions, and cells were counterstained with hematoxylin solution Gill No. 3 (Sigma Aldrich). For the immunofluorescence staining of freshly extracted zfPGCs, dried blood smears were fixed with 4% buffered formaldehyde for 10 minutes at room temperature. Slides were washed 3 times with washing buffer (0.2% Triton X100 in phosphate buffered saline (PBS) and blocked for 1 hour with a blocking solution of 10% pre-immune goat serum in washing buffer. After washing, samples were incubated for 12 hours at 4°C with a primary antibody against stage-specific embryonic antigen 1 (SSEA-1) (1µg/ml; Solter, D./Knowles, B.B.; DSHB Hybridoma Bank) or the primordial germ cell surface marker epithelial membrane antigen 1 (EMA-1) (5µg/ml; Eddy, M./Hahnel, A.; DSHB Hybridoma Bank) diluted in blocking solution. Thereafter, the slides were washed 3 times with washing buffer and incubated for 3 hours at room temperature with the secondary antibody (Alexa Fluor 594 goat-anti mouse (IgM); Life Technologies) diluted 1:300 in blocking solution. Following two washing steps with washing buffer, slides were incubated in 0.1µg/ml DAPI (4′,6-diamidino-2-phenylindole) diluted in water for nuclear staining, washed with PBS and mounted with 50% glycerol in PBS.

For cultured PGCs, the immunostaining was done in suspension with cells that we harvested after 10 days *in vitro* (DIV). Cells were centrifuged for 5 minutes at 2500 g, resuspended in 4% paraformaldehyde, incubated for 10 minutes at room temperature and washed twice with the washing buffer. Subsequently, washed cell pellets were resuspended and incubated in blocking solution for 1 hour at 20°C. After another washing step cells were resuspended in blocking solution containing the primary antibody and incubated for 12 hours at 4°C. Then, the cells were washed again and incubated with the secondary antibody for 3 hours at 20°C. Finally, cells were washed, stained with DAPI, washed again and resuspended in 50% glycerol in PBS. For immunostaining of living PGCs the fixation step was omitted. Images were obtained by performing epifluorescence microscopy.

RT-PCR analyses

Total RNA was isolated from embryonic tissues using the RNeasy mini kit (Qiagen) and from cultured zfPGCs using the Power SYBR® Green Cells-to-Ct^{T} kit (Thermo Fisher), following the manufacturer's instructions. For cDNA synthesis, 2 - 3 µg of total RNA

were denaturated in 10 µl distilled water in the presence of 1µl random hexamers (50 ng/µl) and 1µl desoxyribonukleosidtriphosphates (dNTP mix; 10 mM each) for 5 minutes at 65°C. After chilling the solution on ice for 1 min, we added 2µl 10x reverse-transcriptase (RT) buffer, 4µl 25 mM $MgCl_2$, 2µl 0.1 M 1,4-dithio-D-threitol (DTT), 1µl RNaseOUT™ (40U/µl; Thermo Fisher) and 1µl of the RT SuperScript™ III (200U/µl; Thermo Fisher). Reverse transcription was carried out in a thermocycler for 10 min at 25°C, followed by 50 min at 50°C and, finally, 5 min at 85°C. Next, the reaction mix was placed on ice for 5 min and, after 1µl RNase H (2U/µl; Thermo Fisher) was added, incubated for 20 min at 37°C.

PCR reactions were carried out with the HOT FIREPol® DNA polymerase in buffer B1 (both Solis BioDyne). A 20 µl reaction mixture included 12.5µl H_2O, 2µl 10x buffer B1, 1.6µl $MgCl_2$ (25 mM), 1µl cDNA, 2µl primer pair (10 µM), 0.4 µl dNTP mix and 0.5µl HOT FIREPol®. The thermocycling conditions were as follows: denaturation at 95°C for 10 min, 35 amplification cycles (95°C for 30 sec, 60°C for 40 sec and 72°C for 30 sec) and a final extension at 72 °C for 5 min. The amplified products were resolved by gel electrophoresis in 2% Roti®agarose (Agarose High Resolution; Roth) in 0.5x UltraPure™ TRIS-Borat-EDTA (TBE) buffer (Thermo Fisher). Non-intron spanning primers for DAZL (DAZL(I)), DDX4 (DDX4(I)), NANOG, SOX3 and POU5F1 were the same as in Mak et al. (2015) (Mak et al., 2015). We designed the following intron-spanning primer pairs for zebra finch DAZL (DAZL(II): forward: 5'-GAAACCCAGCACTCAAACGC-3'; reverse: 5'-AAGACGCTCCGAATTTCAGC-3') and DDX4 (DDX4(II): forward: 5'-CTGGAAGCCTACTCCAGTGC-3'; reverse: 5'-TCCCTCATCATTTGGGCCAC-3'). The primer pair designed for ACGT1 was forward: 5'-AACCGGACTGTTTCCAACAC-3'; reverse: 5'-CACCTTCACCGTTCCAGTTT-3'.

Lentiviral transduction and injection of cultured zfPGCs

The lentiviral vector for eGFP under control of the human EF1α-promoter was purchased from SignaGen Laboratories (Gaithersburg MD, USA). All other lentiviral vectors were made in the department of Prof. Dr. Alexander Pfeifer (University of Bonn). For the production of lentiviral vectors, vector plasmids as well as the packaging plasmids pMDLg/pRRE, RSV-rev and pMD2.G were co-transfected into HEK293T cells seeded on poly-L-lysine-coated dishes. The supernatant was collected and centrifuged in an

ultracentrifuge with SW32 Ti rotor at 61,700 g at 17 °C for 2 h. Virus suspensions were concentrated by ultracentrifugation over a 20% (w/v) sucrose cushion in a SW55 Ti rotor at 53,500 g at 17 °C for 2 h. Viral titer was quantified using reverse-transcriptase enzyme-linked immunosorbent assay.

To test for germline transmission, we used the vector for hPGK-eGFP. After 7 DIV, zfPGCs growing in clumps were pooled from 5 - 10 embryos without being separated by sex to reach a density of 1 – 3 x10^3 cells in 50µl of culture medium, and lentiviral vectors were added to a final concentration of 2 x10^8 TU/ml. The culture medium was replaced after 12 hours and the cells were cultured for additional 2 days. To dissociate zfPGC clumps, the cells were washed with PBS 48 hours after viral transduction and digested for 30 minutes at 37°C with a papain (Worthington, LS003119) solution (2mg/ml KnockOut™ DMEM without calcium) that had been sterilized using a 0.2µm syringe filter. Then the zfPGCs were washed with PBS, centrifuged for 5 minutes at 2,500 g and resuspended in culture medium to give a final concentration of 500 cells/µl for injection into a host embryo.

To inject transduced cultured zfPGCs freshly laid zebra finch eggs were incubated for 4 hours at 38 °C and then placed on a silicone surface with the blunt end facing upwards. With a light source that illuminated the egg from below, we were able to observe the blastodisc through the egg shell. Using a scalpel, we opened a 0.5-1 mm window in the egg shell above the blastodisc. Care was taken not to disturb the inner and outer egg membranes, as doing so would reduce the survival of the embryo. Using a pulled glass needle connected to a microinjector, we injected 300 - 500 transduced cultured zfPGCs in the subgerminal cavity of the host embryo. The window in the eggshell of the host embryo was sealed with chicken egg membrane and closed with zebra finch egg shell that had been glued to the host egg shell with zebra finch egg white. After 72 hours of incubation at 38 °C injected eggs that showed host embryo survival and ongoing development were taken to nests of foster parents for further egg incubation, hatching and development until sexual maturity.

Generation and screening of transgenic zebra finches

Founder birds were raised by foster parents until sexual maturity and paired with wildtype

birds. Fertilized eggs produced by the founders were incubated by foster parents that raised the hatchlings until maturity. In order to test for the presence of the eGFP gene in the offspring (F1-birds), blood samples were collected from each hatchling, and genomic DNA (gDNA) was extracted using the DNeasy Blood and Tissue Kit (Qiagen) following manufacturer's instructions to perform PCR analyses. The primers used for the hPGK-eGFP-PCR produced an amplicon (602 bp) that included sequences of both the hPGK promotor and the eGFP gene: the forward primer was 5'-CACTAGTACCCTCGCAGACG-3', the reverse primer was 5'-TCTTGTAGTTGCCGTCGTCC-3'. For each PCR 50 to 100 ng gDNA was used in a 20 µl reaction mixture that contained 12.5 µl H_2O, 2 µl primer mix (10 µM), 0.4 µl dNTPs (10 µM), as well as 2 µl buffer B1 (10X), 1.6 µl $MgCl_2$ (25 mM) and 0.5 µl FIREPol® DNA polymerase (all three from Solis BioDyne). Thermocycling conditions were the following: initial denaturation at 95°C for 5 min, 35 cycles of 95°C for 30 sec, 65°C for 30 sec and 72°C for 30 sec followed by 72°C for 5 min. The amplified products were resolved by electrophoresis in a 2 % Tris-acetate-EDTA (TBE) agarose gel. All steps had been performed in a lab free of eGFP usage or in an UV decontaminated PCR box.

To confirm eGFP expression by RT-PCR the forward primer was 5'-GGGCACAAGCTGGAGTACAA-3' and the revers primer was 5'-AAGGGAGATCCGACTCGTCT-3' For each PCR 50 to 100 ng cDNA was used in a 20 µl reaction mixture that contained 12.5 µl H_2O, 2 µl primer mix (10 µM), 0.4 µl dNTPs (10 µM), as well as 2 µl buffer B1 (10X), 1.6 µl $MgCl_2$ (25 mM) and 0.5 µl FIREPol® DNA polymerase (all three from Solis BioDyne). Thermocycling conditions were the following: initial denaturation at 95°C for 5 min, 40 cycles of 95°C for 30 sec, 68°C for 1 min and 72°C for 10 sec followed by 72°C for 5 min. The amplified products were resolved by electrophoresis in a 2 % Tris-acetate-EDTA (TBE) agarose gel. All steps had been performed in a lab free of eGFP usage or in an UV decontaminated PCR box.

To confirm eGFP expression by immunostaining, F1-birds that were PCR-positive for hPGK-eGFP were sacrificed and transcardially perfused with PBS followed by 4 % paraformaldehyde in PBS. Tissues were post-fixed in 4 % paraformaldehyde, cryoprotected in 10 % and 30 % sucrose in PBS, and sectioned at a freezing microtome. Free-floating tissue sections were washed three times for 10 minutes with PBS then incubated for 2 h in a PBS-blocking solution containing 0.5 % Saponin (Sigma Aldrich 84510) and 10 % pre-immune goat serum (Thermo Fisher). For anti-GFP

immunostaining, sections were incubated for 36 h at 4°C with a chicken anti-GFP antibody (Aves GFP-1020) at a concentration of 1:1000 in PBS-blocking solution. After three subsequent washes with PBS, the samples were incubated for 3 h at room temperature with the goat anti-chicken antibody conjugated with Alexa Fluor® 488 (Abcam 150169) at a concentration of 1:500 in PBS-blocking buffer. Finally, the sections were washed three times with PBS and mounted on a glass slide with VECTASHIELD® (Vector laboratories). Images were obtained by performing epifluorescence microscopy. All birds that were PCR-positive for hPGK-eGFP and GFP-immunopositive were considered to be transgenic (Table 2). Gonads of founders were used freshly for DNA extraction followed by PCR for hPGK-eGFP and after fixation for anti-GFP immunostainings.

To confirm the integration of eGFP in the genome, genome walking was performed using the Universal genomewalker (TAKARA 636406) and following the manufactures instructions.

Separation of cells by density gradient centrifugation

To separate zfPGCs from red blood cells two Ficoll 400 (Sigma F2637) solutions (16 and 6.3 %) were prepared in cell culture medium. Embryonic blood samples were pooled from five individuals and centrifugated for 5 min at 2,500 g. The cell pellet was resuspended in 100 µl of culture medium, mixed gently with 900 µl of the 16 % Ficoll solution, placed underneath 200 µl of the 6.3% Ficoll solution and centrifugated for 30 minutes at 3,400 g. ZfPGCs could be collected from the most superficial layer (200 µl) of the gradient. After 10 µl were removed for cell counting the rest of the cell suspension was diluted with 300 µl of culture medium and centrifugated for 5 min at 2,500 g. After the supernatant was carefully removed the cell pellet was resuspended in the amount of culture medium necessary for RNA extraction.

RNA extraction and next generation sequencing

Four samples with 400 zfPGCs each were collected for each group (1 and 10 DIV) and RNA was extracted using the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing

(TAKARA) following manufactures instructions. Illumina high output sequencing was done by the NGS group at the Max Planck Institute for molecular Genetics.

Statistical analysis

For statistical comparison of zfPGC numbers obtained in cultures with BMP4 and Activin we performed a one-way ANOVA with $p < 0.05$. To analyze the growth of zfPGCs *in vitro* over time we performed a one-way ANOVA followed by a post-hoc test with the p-value being adjusted for multiple comparisons.

The Galaxy web platform (https://usegalaxy.eu/) was used for the analysis of the sequencing data (Afgan et al., 2018). A quality control of the FASTQ files was performed with FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Nextera adapters were removed with Trimmomatic (Bolger et al., 2014). The reads were mapped with STAR (Dobin et al., 2013) and counted with featureCounts (Liao et al., 2014) . A quality control of the mapped sequences was done with MultiQC (Ewels et al., 2016). Differentially expressed genes between the two groups were identified with DESeq2 (Love et al., 2014).

Data Records

The raw data fastq.gz files was deposited in the SRA database under the accession number PRJNA637305.

Acknowledgments

We are grateful to Anja Lohrentz, Christina Reusch, Judith Kammerlander, Antje Bakker, Sunil Nandi and Lorna Taylor for their excellent technical support and Bernd Timmermann for performing the next generation sequencing of our samples. We want to thank David Witkowski, Frances Weigl and Frank Lehmann for their outstanding care of the birds in this study. Thanks go to Luisana Carballo for comments on the manuscript draft. IG also thanks the IMPRS research school for training and support. The study was founded by the Max Planck Society.

Supplemental information

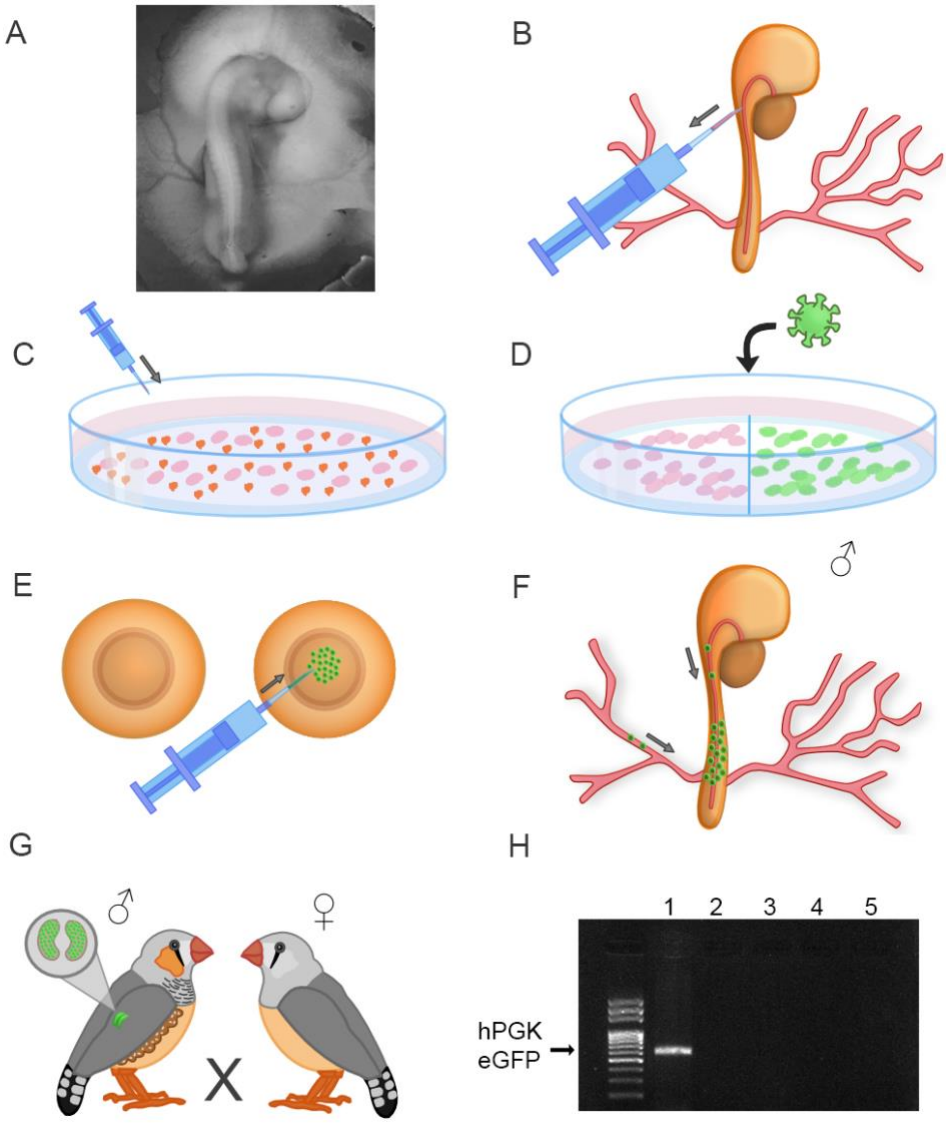


Figure S1. Method of transgenic zebra finch generation using cultured primordial germ cells. Related to Materials and methods

(A) Microphotograph of a zebra finch embryo at Murray stage 15. (B-G) Schematic representation of the procedure: **(B)** Blood extraction from the dorsal aorta of zebra finch embryos at Murray stage 13 to 15. **(C)** Selected expansion of zebra finch primordial germ cells (zfPGCs) *in vitro*. While erythrocytes started to die after 4 days *in vitro* (DIV), the proliferation of zfPGCs continued. **(D)** After 7 DIV, zfPGCs were transduced with lentiviral vectors for enhanced green fluorescent protein (eGFP) under control of the human phosphoglycerate kinase (hPGK) promoter. **(E)** zfPGCs that expressed eGFP at 10 DIV were harvested and injected into host embryos at the blastodisc stage. **(F)** During incubation of the injected eggs at 38°C for 60 hours eGFP-positive zfPGCs migrated through the developing circulatory system to the gonadal *anlage*. **(G)** Zebra finches with gonads that enclosed eGFP-expressing zfPGCs were raised to maturity and mated with wildtype birds. **(H)** The presence of the eGFP gene was determined by PCR of blood samples taken from the founders' offspring. In blood samples that were taken from birds of the same clutch (1-5), transgenic offspring was identified by the presence of the hPGK-eGFP band (1).

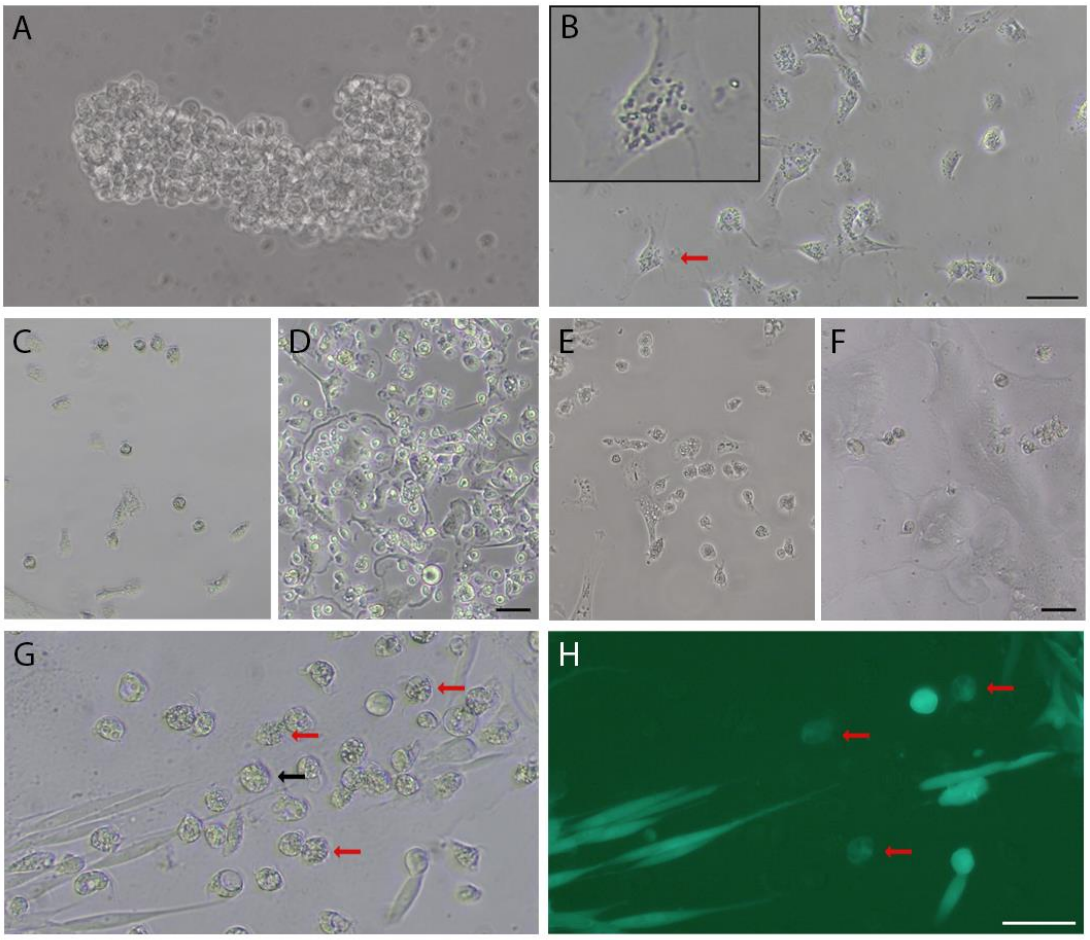

A
B
C
D
E
F
G
H

Figure S2. Comparisons of culture conditions for zfPGCs extracted from embryonic blood and gonads. Related to Figure 1

Phase contrast images are shown for embryonic blood-derived zfPGCs cultured for 10 DIV in **(A)** the adapted FAIcs culture medium (Tab. S1) and in **(B)** gonadal PGC culture medium (Jung et al., 2019). The insert in (**B**) shows a zfPGC (red arrow) at higher magnification. Note that embryonic blood-derived zfPGCs cultured in adapted FAIcs conserved a rounded shape when growing in clumps (**A**) but attached to the bottom of the cell culture plate, changed their morphology and died shortly after when cultured in the gonadal PGC culture medium (**B**). In **(C - F)** phase contrast images are presented for zfPGCs that were extracted from embryonic gonads (zfgPGCs) and cultured in gonadal PGC culture medium (**C, D**) and adapted FAIcs (**E, F**), respectively. Images were taken after 1 (**C, E**) and 10 (**D, F**) DIV, respectively. Note that after 10 DIV in gonadal PGC culture medium zfgPGCs have proliferated and a feeder layer was formed with stromal cells (**D**) whereas in adapted FAIcs the zfgPGCs died (**F**). However, zfgPGCs were transduced with a lentiviral vector at relative low efficiency even when cultured in gonadal PGC culture medium. Phase contrast and **(G)** fluorescent **(H)** images of zfgPCRs are shown after transduction with a lentiviral vector for eGFP under control of the human UBC (hUBC) promoter. In (**G, H**) red arrows point to successfully transduced zfGPGCs and in (**G**) the black arrow hint at one of the zfGPGCs that did not express eGFP.

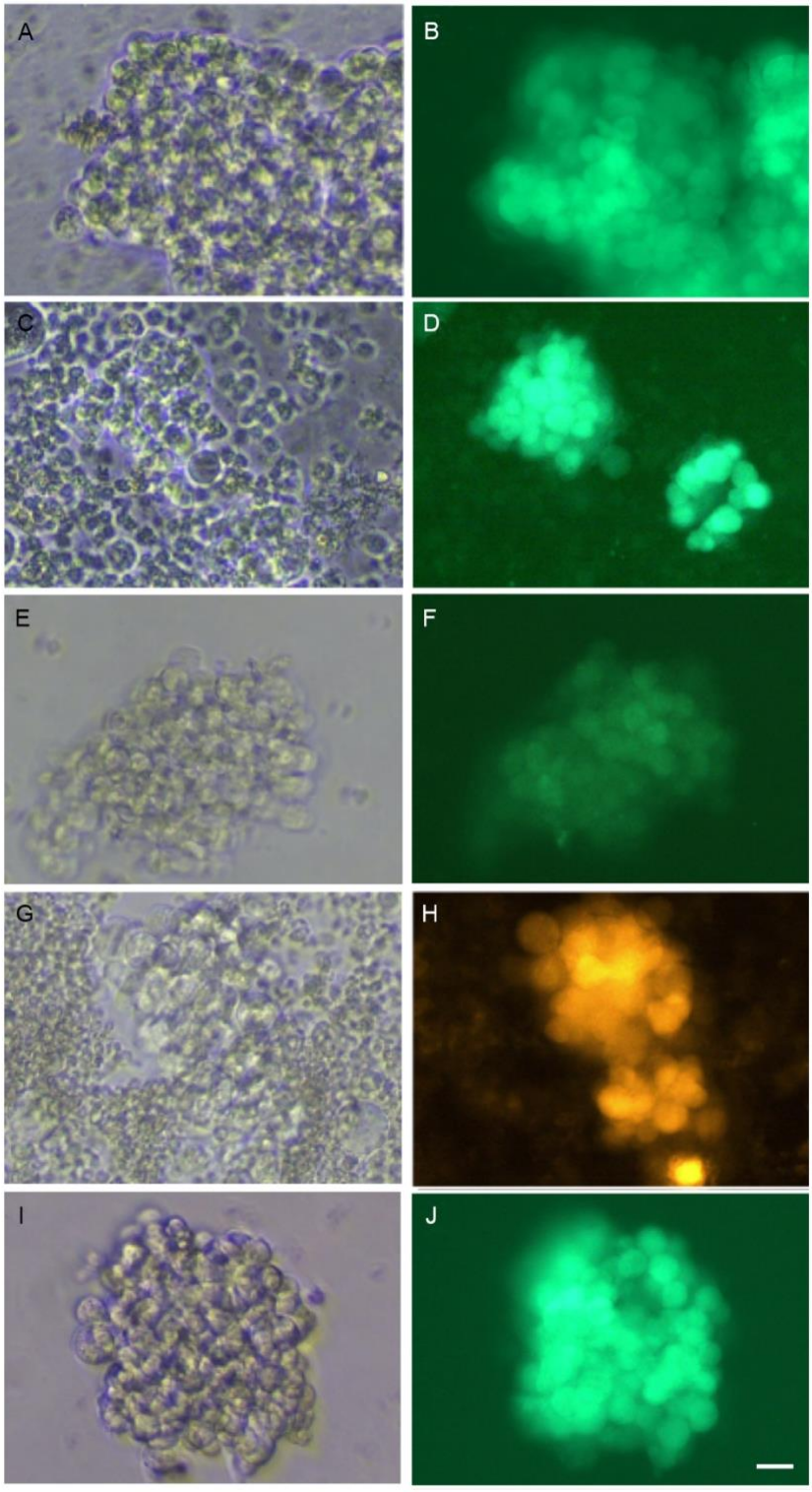

A
B
C
D
E
F
G
H
I
J

Figure S3. Lentiviral vectors containing different promoters were capable of driving the reporter gene expression in cultured zfPGCs. Related to Figure 3

Phase contrast (left panels) and fluorescent (right panels) images of cultured zfPGC clumps that were transduced with lentiviral vectors for eGFP (**A - D, I, J**) and tomato (**E, F**). Lentiviral constructs contained different promoters: **(A, B)** human phosphoglycerate kinase (hPGK) promoter, **(C, D)** human ubiquitin C (hUBC) promoter, **(E, F)** human elongation factor 1 alpha (hEF1α) promoter, **(G, H)** cytomegalus virus (CMV) promoter and **(I, J)** CMV enhancer fused to the chicken beta-actin (CAG) promoter. The titer we used was 2x10exp8 TU/ml. Scale bar represents 20 μm.

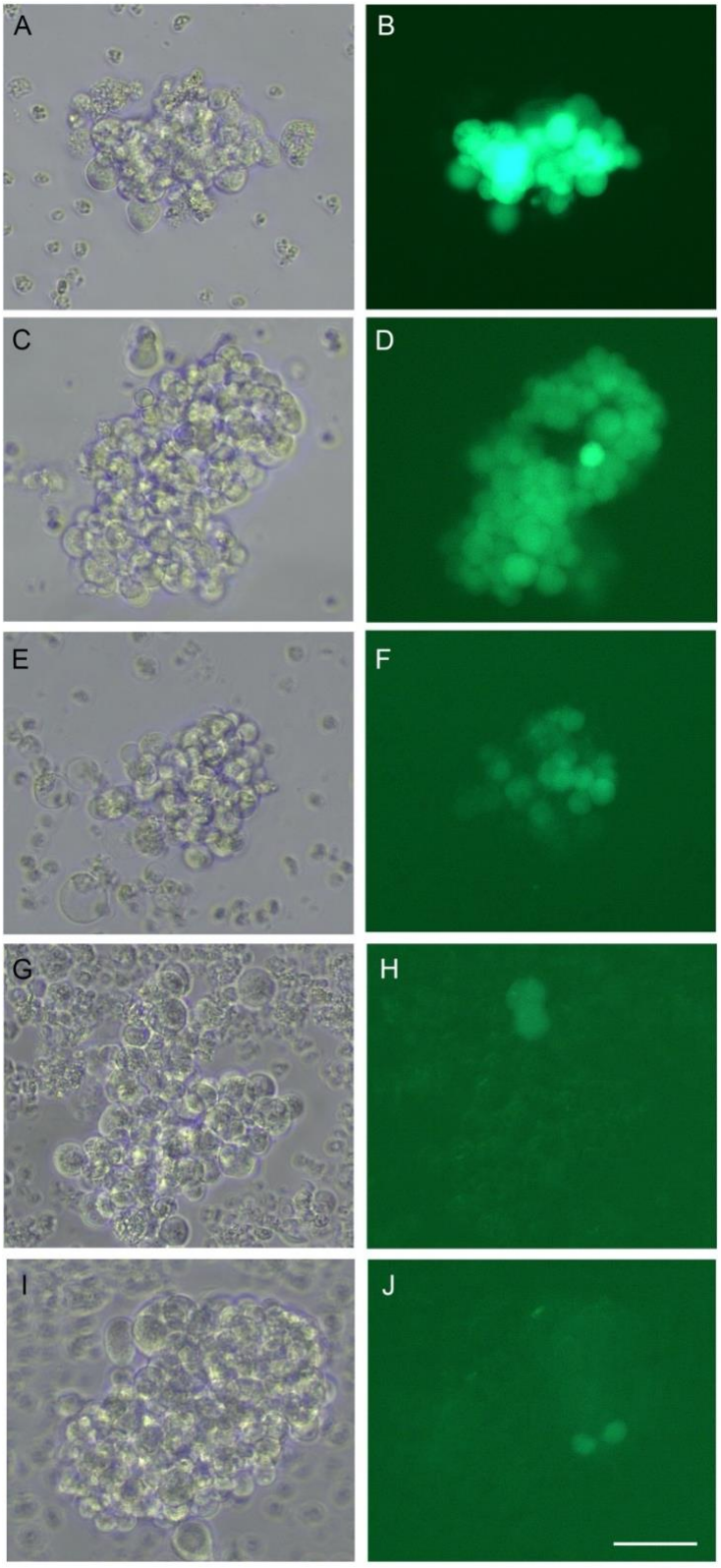

A
B
C
D
E
F
G
H
I
J

Figure S4. Determination of effective lentiviral titers for the transduction of cultured zfPGCs. Related to Figure 3

Phase contrast (left panels) and fluorescent (right panels) images of cultured zfPGC clumps that were transduced with the lentiviral vector rrl-hPGK-eGFP using different viral titers (final concentrations): 4x10exp8/ml TU/ml **(A, B)**, 2x10exp8 TU/ml **(C, D)**, 4x10exp7 TU/ml **(E, F)**, 2.7x10exp7 TU/ml **(G, H)** and 8x10exp6 TU/ml **(I, J)**. Note strong expression of eGFP at a concentration of 2x10exp8 TU/ml and above. Scale bar represents 50 μm.

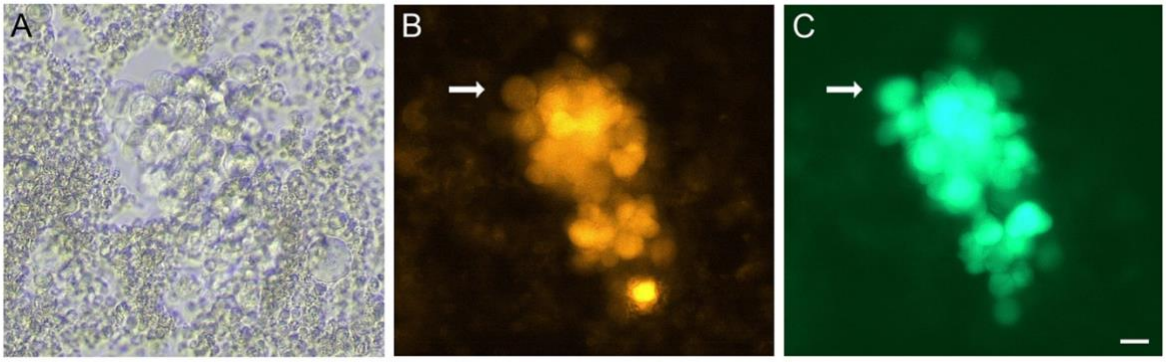

A
B
C

Figure S5. Transduction of cultured zfPGCs with two different lentiviral vectors. Related to Figure 3

Microphotographs of cultured zfPGC clumps that were simultaneously transduced with lentiviral vectors for hUBC-eGFP and CMV-tomato. Phase contrast image **(A)** as well as fluorescent images of eGFP **(B)** and tomato **(C)**, are shown. White arrows point to the same zfPGC being fluorescent with both filters. Note that most zfPGCs expressed both eGFP and tomato. Scale bar represents 20 μm.

Supplemental Table 1. FAIcs culture medium adapted to culture zfPGCs. Related to Materials and Methods

Culture medium components		Concentration	Brand	Catalog number
Basal medium	KnockOut^{TM} DMEM without Calcium	70%	Thermo Fischer Scientific (TFS)	Custom made
	KnockOut^{TM} DMEM with Calcium	30%	TFS	10829018
Supplements	B27 Supplement	1,5%	TFS	17504044
	Ovalbumin	1%	Sigma Aldrich	A-5503
	GlutaMax	1%	TFS	35050061
	Non-Essential Amino Acid	1%	TFS	11140-050
	Nucleosides	1%	Sigma Aldrich	ES-008-D
	Fetal Bovine Serum	0.4%	Sigma Aldrich	12103-C
	Sodium Heparin	0.2%	Sigma Aldrich	H3149
	β-Mercaptoethanol	0.25 mM	TFS	31350-010
	Pyruvate	0.2 mM	TFS	11360-039
	Cholesterol	4000 ng/ml	Sigma Aldrich	C3045-5G
	Chicken BMP4	25 ng/ml	LSBio	LS-G11280
	Human FGF2	5ng/ml	RandD systems	234-FSE-025
	Human IGF	25ng/ml	RandD systems	291-G1-200

Supplemental Table 2 Incidence of GFP-immunopositive cells in transgenic birds. Related to Figure 4

Organs	Transgenic birds with immunopositive cells (%)	Abundance of immunopositive cells (%)
Liver	100	1-20
Kidney	100	4-51
Brain	92	22-70
Heart	92	12-43
Lung	62	3-39
Ovary	60	15-29

Chapter II

Identification and evaluation of cell culture conditions that support the proliferation of zfPGCs

Ivana Gessara

Falk Dittrich

Moritz Hertel

Carolina Frankl-Vilches

Mike McGrew

Manfred Gahr

Manuscript in preparation for publication

Identification and evaluation of cell culture conditions that support the proliferation of zfPGCs

Authors

Ivana Gessara[1]*, Falk Dittrich[1], Moritz Hertel[1], Carolina Frankl-Vilches[1], Mike McGrew[2], Manfred Gahr [1]

1. Max Planck Institute for Ornithology, Eberhard Gwinner Strasse, 82319 Seewiesen, Germany

2. The Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian EH25 9RG, United Kingdom

Abstract

Primordial germ cells (PGCs), the precursors of oocytes and sperm cells, can be used in avian species to generate transgenic animal models. As part of this process, PGCs are cultured *in vitro*, but long-term cultures have only been established for chicken PGCs so far. In songbirds PGCs can be cultured from the zebra finch (zfPGC), exclusively. However, zfPGCs can only be culture for a relative short period of time that does not allow for the selection and subsequent proliferation of genetically modified PGCs. Since we detected the growth curve of zfPGCs in culture to reach the stationary phase between 10 and 15 days *in vitro* (DIV) we performed a comparative transcriptome study of zfPGCs cultured for 1 and 10 DIV. The goal was to determine the signaling pathways affected in zfPGCs during culturing and to identify growth factors necessary to maintain long-term zfPGC self-renewal. We found the expression of genes associated with cell survival and cell proliferation signaling pathways PI3K/AKT, MEK/ERK and TGF-beta to be affected in zfPGCs after 10 DIV. Selecting the respective growth factors Vascular Endothelial Growth Factor (VEGF), Leukemia inhibitory factor (LIF), Epidermal Growth Factor (EGF), Stem Cell Factor (SCF) and Transforming Growth Factor beta (TGF-ß) to stimulate these pathways in zfPGC cultures, we observed a significant effect in the proliferation of zfPGCs when cultured in the presence of VEGF and LIF, as well as a positive trend when cultured in the presence of EGF. VEGF and LIF did not increase the survival time of zfPGCs *in vitro* but contributed to obtain a larger number of zfPGCs in a short period of time in culture.

Introduction

Primordial germ cells (PGCs) are the precursors of spermatozoa and oocytes, the only cell types capable of passing the genetic information to the next generation. PGCs are used in avian species to generate transgenic birds. They can be extracted from the bloodstream of Hamburger Hamilton (HH) stage 14 HH embryos, expanded in culture, genetically modified *in vitro* and reinjected into Eyal-Giladi and Kochav (EGK) stage X or HH stage 14 embryos. Once injected into such early embryos, the exogenous PGCs will migrate to the gonadal anlage, contribute to the germline and ultimately lead to the generation of transgenic offspring (Van De Lavoir et al., 2006; MacDonald et al., 2010; Gessara et al., 2021).

The long-term culture of PGCs is of great importance when applying precise genome editing techniques such as the CRISPR-Cas with homology directed repair method. To apply this gene editing approach, it seems to be necessary to start with a larger number of cells that can be genetically modified and to be able to expand clonally genetically altered cells for an extended period in culture. So far, the chicken is the only bird species in which PGCs can be expanded in culture for an unlimited period of time. This is possible even without a cellular feeder layer and in the presence of very low serum concentration in the culture medium (Whyte et al., 2015).

Recently, two culture methods were developed to expand zebra finch PGCs *in vitro* for the first time and used for the generation of transgenic zebra finch models (Jung et al., 2019; Gessara et al., 2021). The first one proposes the use of zfPGCs obtained from embryonic gonads and requires a high serum concentration in the medium as well as a cellular feeder layer made of stromal gonadal cells (Jung et al., 2019). The second one we developed to expand *in vitro* zfPGCs obtained from embryonic blood using low concentrations of serum and relinquishing a cellular feeder layer for the zfPGCs to grow *in vitro* (Gessara et al., 2021). Both methods are efficient enough to maintain zfPGCs proliferation *in vitro* for two weeks to a month. The culture mediums used in those studies include the growth factors FGF2 and BMP4 that are known to stimulate three different signaling pathways related to cell proliferation and cell survival MEK/ERK, PI3K/AKT and TGF-ß (Böttcher and Niehrs 2005; De Felici 2000; Wang et.al 2014; Pauklin and Vallier 2015). These pathways are stimulated by the interaction of specific growth factors such as Fibroblast Growth Factor (FGF), Insulin-like Growth factor (IGF), Bone Morphogenetic Protein (BMP), Activin A and Stem Cell Factor (SCF) among others, with their respective receptors. In particular, binding of FGF to the FGF receptor leads to the autophosphorylation of the kinase domains. The phosphorylated FGF receptor can bind to the Growth Factor Receptor Bound Protein 2 (GRB2) which interacts with Son of Sevenless (SOS) that activates the Ras GTPase and subsequently the RAF kinases. After Raf activation the phosphorylation of the kinase of the MSAP-kinase (MEK) is induced and this event activates the Extracellular-signal Regulated Kinases (ERK) cascade inducing cell proliferation (Figure 1 A). FGF can also activate the PI3K/AKT cascade via Ras or interacting with GRB2 associated Binding Protein 1 (GAB1) that binds to GRB2 and phosphorylates and activates the Phosphatidylinositol-4,5-Bisphosphate-3 Kinases (PI3K). PI3K triggers Phosphatidylinositol-4,5-Bisphosphate (PIP2) to generate Phosphatidylinositol-3,4,5-Trisphosphate (PIP3) activating the AKT

Serine/Threonine kinases (AKT) through phosphorylation; AKT will then promote cell survival and protein synthesis and will inhibit apoptosis (Stoica et al. 2003; Böttcher and Niehrs 2005) (Figure 1 A). IGF and SCF can play a similar role as FGF: IGF when interacting with the IGF receptor activating the MEK/ERK cascade via GRB2 or when activating the Insulin Receptor Substrate (IRS) which can also initiate the PI3K/AKT cascade (Singh et al., 2013) (Figure 1 B). SCF when interacting with its C-Kit tyrosine kinases receptor activates the PI3K/AKT cascade directly and the MEK/ERK cascade through GRB2 leading to the stimulation of cell proliferation and survival (Sette et al. 2000; De Felici 2000; Liu et al. 2007) (Figure 1C) . Lastly, the TGF-ß signaling pathway can be induced by TGF-ß, BMP4 and Activin A when binding to their receptors which phosphorylate SMAD 1/5/8 (BMP4) or SMAD 2/3 (TGF-ß and Activin). When associate with SMAD 4 the SMAD proteins translocate into the nucleus and promote cell proliferation (Wang et.al 2014; Pauklin and Vallier 2015) (Figure 2).

These growth factors when used in mouse and chicken PGC cultures helped to maintain the expansion of PGCs *in vitro* in the long term (Buehr 1997; Van de Lavoir et al. 2006; Dolci et al. 1991; Godin et al. 1991; Matsui et al. 1991; Manova et al. 1992; Pesce et al. 1993; Karagenc and Petitte, 2000; Whyte et al., 2015).

The aim of our transcriptome study was first to determine changes in the gene expression of zfPGCs after being exposed to culture conditions that could be responsible for a reduced ability of zfPGCs to proliferate and survive *in vitro.* And second more specifically to detect changes in the expression of genes that are part of the MEK/ERK, PI3K/AKT and TGF-ß signaling pathways due to culture conditions. The effect of incorporating respective candidate growth factors to the culture medium on zfPGC self-renewal *in vitro was* determined.

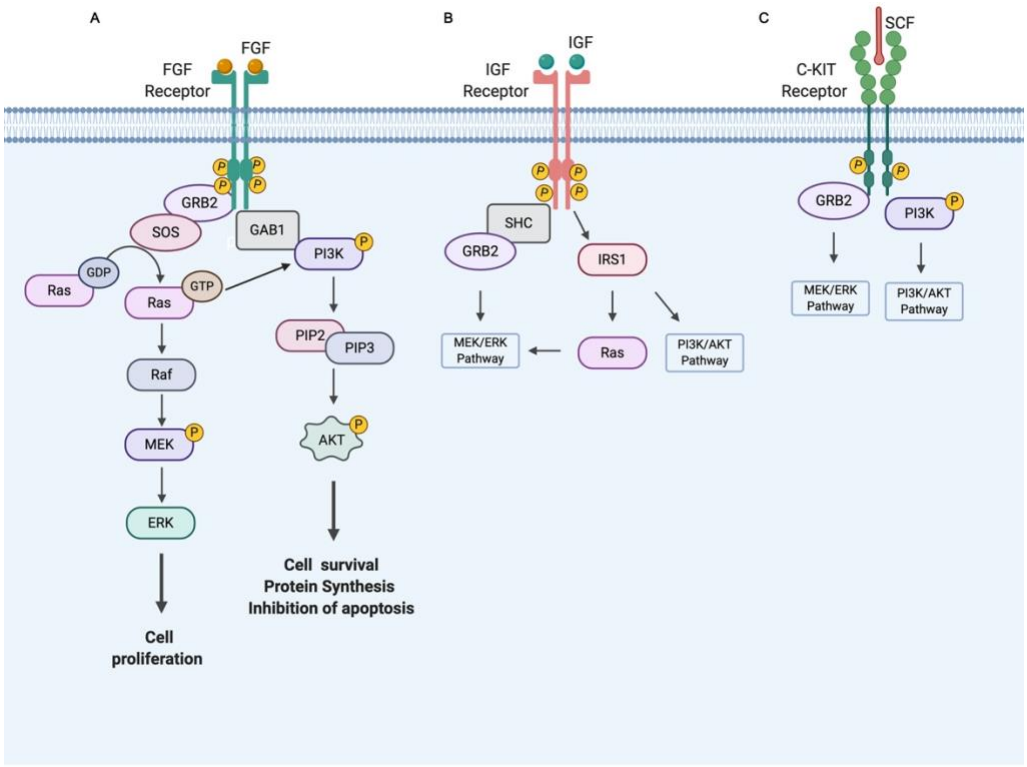


Figure 1. MEK/ERK and PI3K/AKT signaling pathways. The MEK/ERK and PI3K/AKT pathways can be activated when growth factors such as FGF (A) IGF (B) or SCF (C) bind to their receptors stimulating the activation of transcription factors and the consequent expression of genes related to cell proliferation and cell survival.

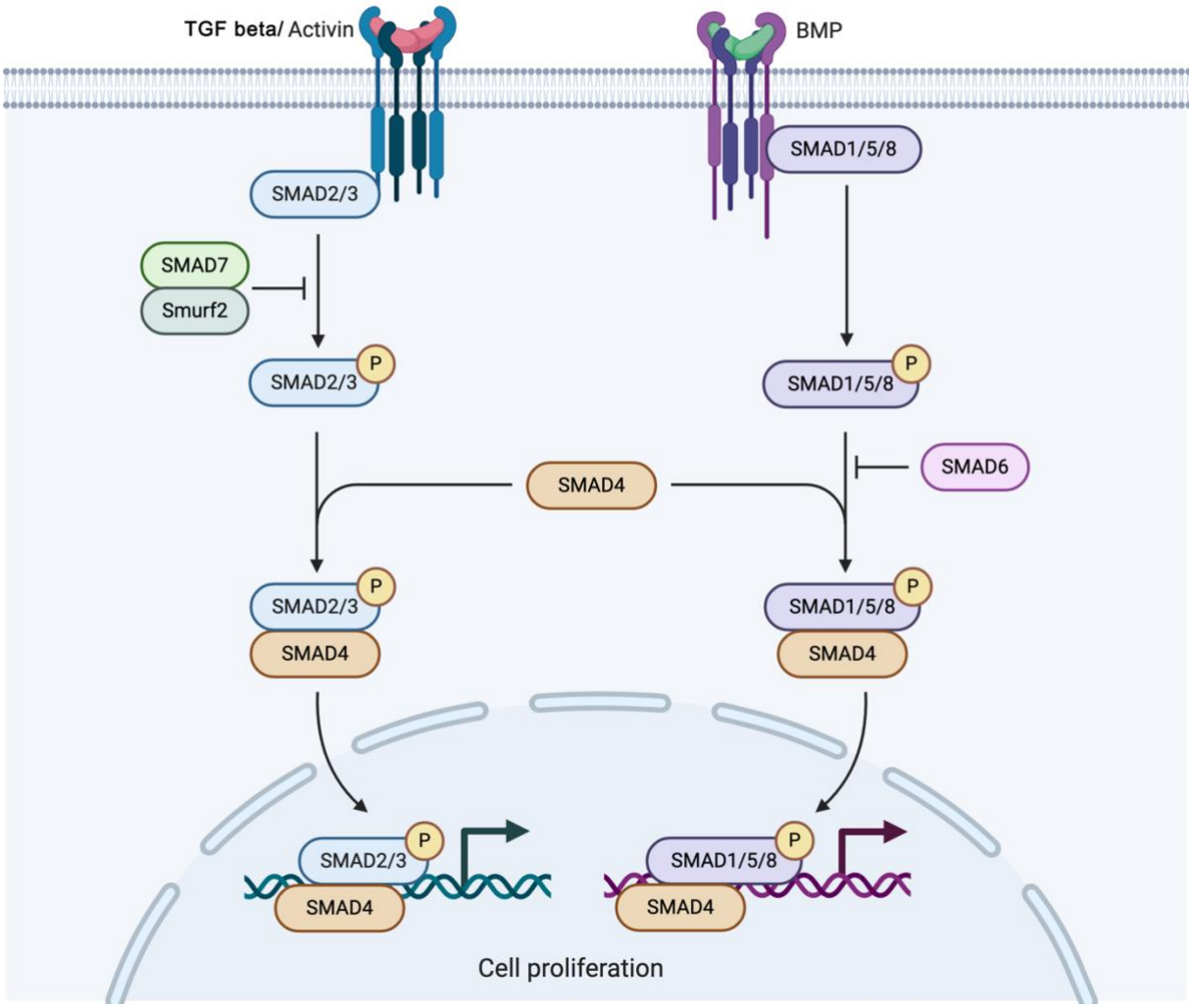


Figure 2. TGF-ß signaling pathway. The TGF-ß signaling pathway is activated when growth factors such as TGF-ß, Activin and BMP4 interact with their receptors. The interaction between SMAD 2/3 (in the case of TGF-ß or Activin) or SMAD 1/5/8 proteins (in the case of BMP) with SMAD 4 lead to the translocation into the nucleus and the activation of genes related to cell proliferation. SMAD 6 and 7 and Smurf2 act as inhibitors of the pathway.

Results

Transcriptome analysis of cultured zfPGCs revealed an increase of the oxidoreductase activity after 10 DIV

To compare the gene expression profile of cultured zfPGCs being in different time phases of the growth curve *in vitro*, for each group we collected four samples of zfPGCs that were cultured for one day (group 1; log phase) and ten days (group 2; stationary phase). As expected, we observed a similar gene expression pattern between samples that belonged to the same group and a clear separation between samples that belonged to different groups (Figure 3A). Setting the cutoff at a 1.5-fold change and $P < 0.05$ we found 2430 genes to be down- and 2617 genes to be up-regulated in zfPGCs obtained after 10 DIV. Up- and down-regulated genes showed a fairly even dispersion of the data points in the MA-plot (Figure 3B) and a balanced volcano plot (Figure 3C). Using the transcriptomic data, we determined the gene ontology groups and pathways related to biological processes that were affected by the period in culture. We found 20 groups significantly overrepresented in group 1, hence downregulated in group 2 (Figure 4 A). Strikingly, we found the group oxidoreductase activity acting on peroxidase as acceptor to be the most affected group, with 33% of the pathways component of the group and 50 genes to be downregulated in group 2 (Figure 4 A-C). Since these gene expression changes could increase the oxidative stress and promote apoptosis, we tested the long-term effect of antioxidants on zfPGCs *in vitro*. At 20 DIV we could not find a significant difference in the proliferation rate between zfPGCs cultured in the presence and absence (control) of antioxidants (Figure 4 D).

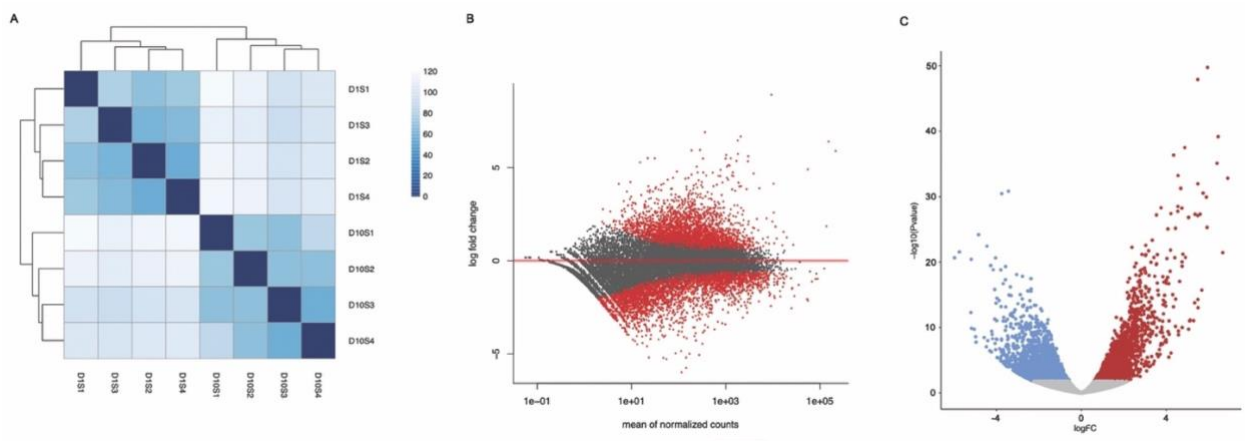
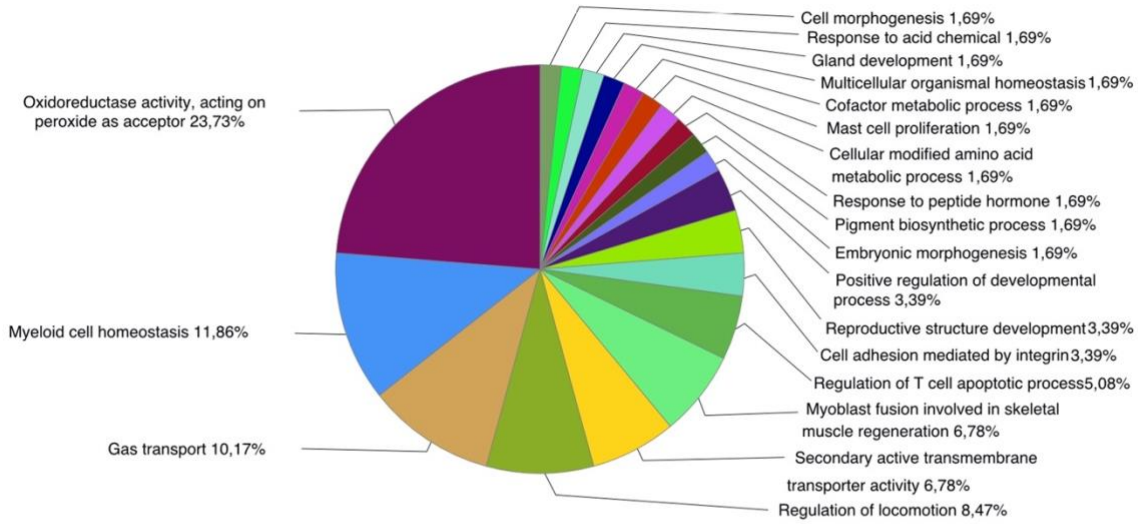


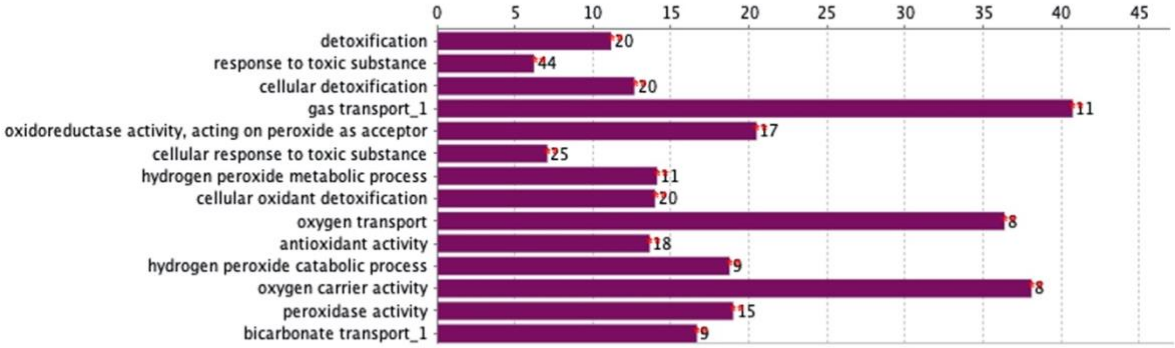
Figure 3. Differential gene expression between zfPGCs cultured for one and ten DIV. (A) Expression heat map of sample-to-sample distances on the matrix of variance-stabilized data for overall gene expression. Darker colors indicate more similar expression. Clustering demonstrates that the replicates from zfPGCs with 1 DIV (D1S1-;

group 1) are very similar to each other but completely separated from replicates of zfPGCs with 10 DIV (D10S1-4; group 2). (B) MA-plot showing gene expression levels in group 1 compared to group 2. The red line marks the point where there is no difference between the two groups. (C) Volcano-plot showing up- (red data points) and down-regulated (blue data points) genes of group 1. Data points in gray represent genes exhibiting non-significant expression changes ($P > 0.05$).

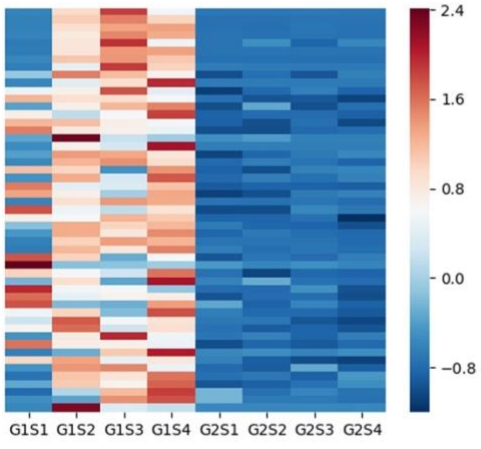
A


Oxidoreductase activity, acting on peroxide as acceptor 23,73%
Myeloid cell homeostasis 11,86%
Gas transport 10,17%
Cell morphogenesis 1,69%
Response to acid chemical 1,69%
Gland development 1,69%
Multicellular organismal homeostasis 1,69%
Cofactor metabolic process 1,69%
Mast cell proliferation 1,69%
Cellular modified amino acid metabolic process 1,69%
Response to peptide hormone 1,69%
Pigment biosynthetic process 1,69%
Embryonic morphogenesis 1,69%
Positive regulation of developmental process 3,39%
Reproductive structure development 3,39%
Cell adhesion mediated by integrin 3,39%
Regulation of T cell apoptotic process 5,08%
Myoblast fusion involved in skeletal muscle regeneration 6,78%
Secondary active transmembrane transporter activity 6,78%
Regulation of locomotion 8,47%

B


0 5 10 15 20 25 30 35 40 45
detoxification 20
response to toxic substance 44
cellular detoxification 20
gas transport_1 11
oxidoreductase activity, acting on peroxide as acceptor 17
cellular response to toxic substance 25
hydrogen peroxide metabolic process 11
cellular oxidant detoxification 20
oxygen transport 8
antioxidant activity 18
hydrogen peroxide catabolic process 9
oxygen carrier activity 8
peroxidase activity 15
bicarbonate transport_1 9

C


2.4
1.6
0.8
0.0
−0.8
G1S1 G1S2 G1S3 G1S4 G2S1 G2S2 G2S3 G2S4

D

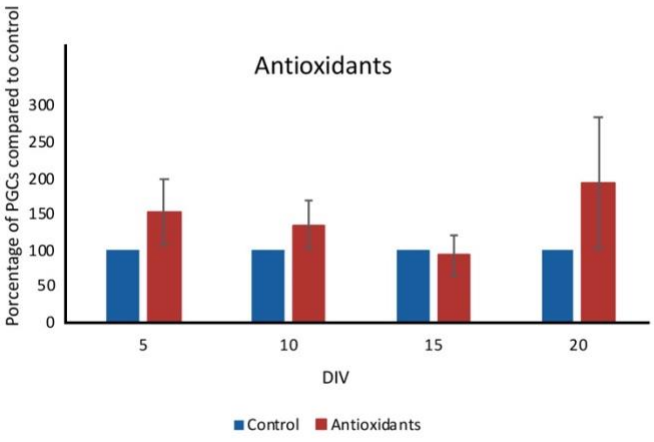

Antioxidants
Porcentage of PGCs compared to control
300
250
200
150
100
50
0
5
10
15
20
DIV
Control
Antioxidants

Figure 4 Gene ontology analysis revealed antioxidant genes to be down-regulated in zfPGCs at 10 DIV. (A) Gene ontology groups overrepresented in zfPGCs cultured for 1 DIV. (B) pathways component of the gene ontology group oxidoreductase activity acting on peroxidase as acceptor, numbers at the top indicate percentage of genes present per pathway, numbers at the end of bars indicate genes upregulated in each pathway, *P < 0.05. (C) Gene expression heatmap of genes with antioxidant activity in four zfPGC samples obtained at 1 (G1S1-S4) and 10 (G2S1-S4) DIV. Z scores obtained from read counts are indicated by different colors in a range from blue (low) to red (high). (D) Effect of antioxidants (Sigma A1345) on the growth of zfPGCs *in vitro*. Blood samples from single embryos were split in two and cultured in the presence or absence of the antioxidant supplement. Cell numbers were counted every 5 days, and the mean percentage of cells in relation to the controls are given (mean ± SEM; N=5).

Key apoptotic proteins remained unaltered or down-regulated in zfPGCs at 10 DIV.

The downregulation of pathways related to oxidoreductase activity acting on peroxidase as acceptor might increase the concentration of hydrogen peroxide in the cytoplasm of zfPGCs leading to DNA damage, telomere shortening, G1 cell cycle arrest and finally apoptosis (Duan et al., 2005).

The intrinsic and extrinsic apoptotic pathways are initiated by the cleavage of caspase-3 and result in DNA fragmentation, degradation of cytoskeletal and nuclear proteins and formation of apoptotic bodies (Elmore 2007). Caspases can be categorized into initiators (caspase 2, 8, 9, 10), effectors (caspase 3, 6, 7) and inflammatory caspases (caspase 1, 4, 5) (Cohen, 1997; Rai et al., 2005). In order to determine if the cause of the stall in proliferation in zfPGCs after 10 DIV was related to the telomere shortening or activation of apoptosis pathways, we analyzed the expression of telomerase reverse transcriptase (TERT) and caspases in zfPGCs cultured for 1 and 10 DIV. TERT and Caspases 2, 3, 8, 9 and 10 were found to be equally expressed in zfPGCs cultured for 1 and 10 DIV while caspases 6 (log2 fold change -1.81) and 7 (log2 fold change -2.68) where found to be down-regulated in zfPGCs cultured for 10 DIV. These expression pattern could indicate that telomere shortening and apoptosis might not be the cause of proliferation arrest in zfPGCs cultured for 10 DIV. For alternative effects of hydrogen peroxide including cell

cycle arrest (Santa-Gonzalez et al., 2016), the transcriptome data did not provide information to further analyze this possibility.

MEK/ERK, PI3K/AKT and TGF-ß signaling pathways in zfPGC at 10 DIV

The standard culture conditions for zfPGCs include the growth factors FGF2 and BMP4 (Jung et al., 2019; Gessara et al, 2021). The MEK/ERK and PI3K/AKT pathways were found to be partially up-regulated in zfPGCs cultured for 10 DIV, presumably due to the presence of FGF2 in the medium (Supplemental Table 1, 2). However, some of the components of the MEK/ERK pathway such as SHC, MEK1, MEK3, MEK4 and ERK3 were found to be down-regulated after 10 DIV (Supplemental Table 1, 2). This gene expression changes could indicate a compensatory response due to the high concentration of FGF2 in the medium.

The TGF-ß signaling pathways though SMAD 1,5 and 8 was found to be down-regulated in cells cultured for 10 days and the pathway inhibitor SMAD 6 was up-regulated in cells cultured for 10 days (Supplemental Table 3). This suggests a lack of response to human BMP4.

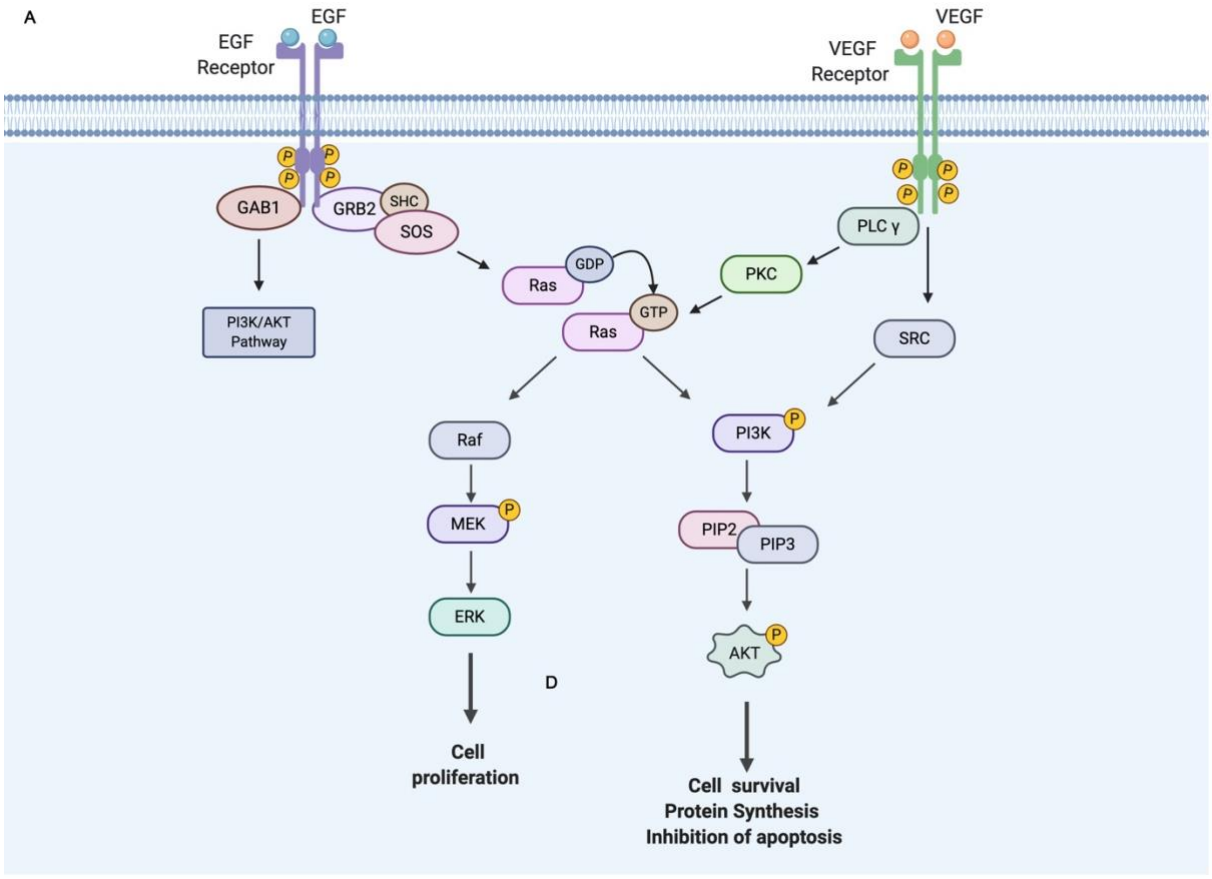
VEGF stimulated zfPGC proliferation in vitro

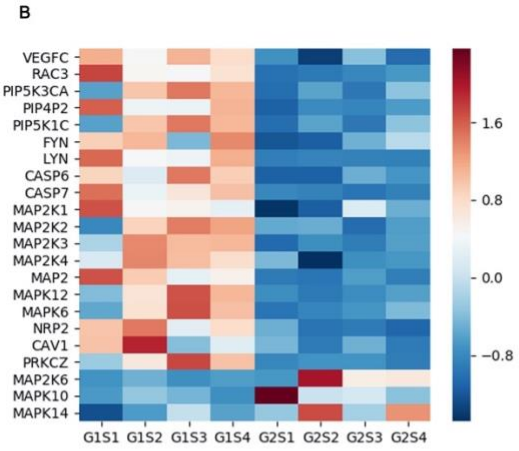
VEGF and EGF stimulate cell survival, proliferation and migration through the MEK/ERK and PI3K/AKT pathways (Matsumoto et al., 2001; Scaltriti et al., 2006). VEGF binding to the VEGF receptors activates the Phospholipase C Gamma 1 (PLCG1) and in consequence the protein Kinase C (PKC) which initiates the MEK/ERK cascade through RAS leading to cell proliferation. Activated VEGF receptors also activate the SRC-Kinase which initiates the PI3K/AKT leading to cell survival (Figure 5 A).

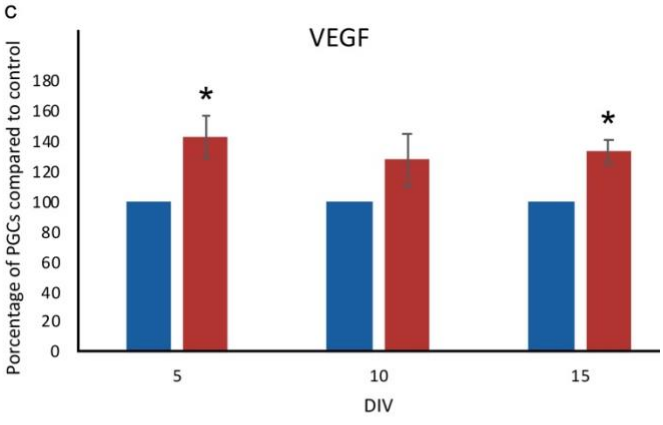
EGF binds to the EGF receptor and activates the MEK/ERK signaling cascade though the activation of GRB2 and SOS with or without the adaptor molecule SHC and the PI3K/AKT signaling cascade though GAB1 or RAS (Figure 5 A). For the signaling pathways activated by EGF see Figure 5A.

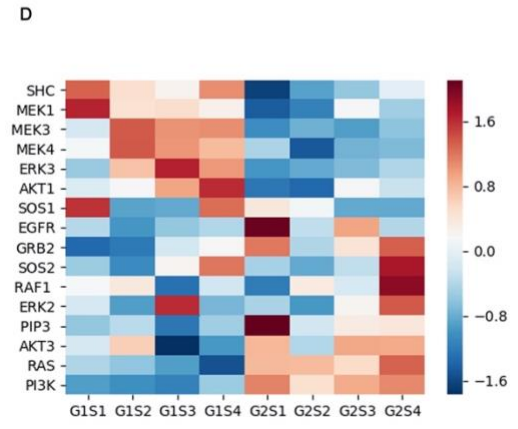
We found several genes that are part of these pathways to be down-regulated in zfPGCs at 10 DIV (Figure 5 B, D). For this reason, we selected these growth factors to test their effect on zfPGC proliferation *in vitro*. When we cultured zfPGCs in the presence of VEGF and EGF for 15 DIV, we found a significantly higher percentage of cells in

relation to the control culture in cultures containing VEGF P=0.001 N=10, and a positive trend in cultures containing EGF P=0.1 N= 4 (Figure 5 C, E).


A
EGF
EGF Receptor
VEGF
VEGF Receptor
GAB1
GRB2
SHC
SOS
PLC γ
PKC
Ras
GDP
GTP
Ras
SRC
PI3K/AKT Pathway
Raf
PI3K
MEK
PIP2
PIP3
ERK
AKT
D
Cell proliferation
Cell survival
Protein Synthesis
Inhibition of apoptosis


B
VEGFC
RAC3
PIP5K3CA
PIP4P2
PIP5K1C
FYN
LYN
CASP6
CASP7
MAP2K1
MAP2K2
MAP2K3
MAP2K4
MAP2
MAPK12
MAPK6
NRP2
CAV1
PRKCZ
MAP2K6
MAPK10
MAPK14
G1S1 G1S2 G1S3 G1S4 G2S1 G2S2 G2S3 G2S4
1.6
0.8
0.0
−0.8


C
VEGF
Porcentage of PGCs compared to control
0
20
40
60
80
100
120
140
160
180
*
*
5
10
15
DIV


D
SHC
MEK1
MEK3
MEK4
ERK3
AKT1
SOS1
EGFR
GRB2
SOS2
RAF1
ERK2
PIP3
AKT3
RAS
PI3K
G1S1 G1S2 G1S3 G1S4 G2S1 G2S2 G2S3 G2S4
1.6
0.8
0.0
−0.8
−1.6

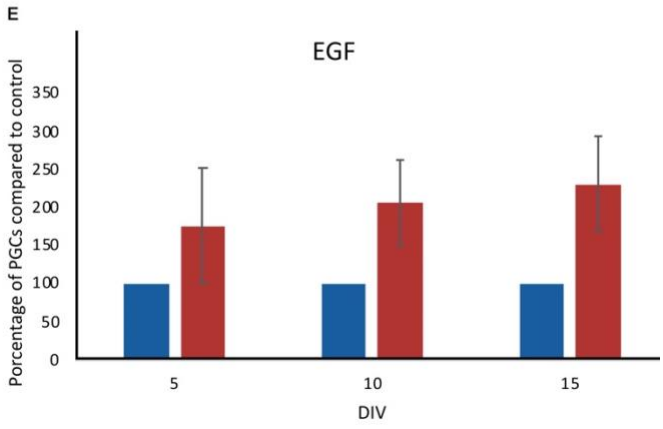

E
EGF
Porcentage of PGCs compared to control
0
50
100
150
200
250
300
350
5
10
15
DIV

Figure 5 VEGF and EGF: signaling pathways and their effect on zfPGC growth *in vitro*. (A) EGF and VEGF signaling pathways. (B,D) Gene expression heatmaps of genes part of VEGF (B) and EGF (D) signaling pathways in four zfPGC samples obtained at 1 (G1S1-S4) and 10 (G2S1-S4) DIV. Z scores obtained from read counts are indicated by different colors in a range from blue (low) to red (high). (C, E) Effect of VEGF (C) and EGF (E) on the growth of zfPGCs *in vitro*. Blood samples from single embryos were split in two and cultured in the presence or absence of VEGF (C) or EGF (E). Cell numbers were counted every 5 days and the mean percentage of cells in relation to control are given, (mean ± SEM; *P < 0.05; N=4 for EGF and N=10 for VEGF).

LIF significantly increased the cell number in zfPGCs cultured for 15 days

LIF participates in two main process, the maintenance of the undifferentiated state of somatic and germinal stem cells and the stimulation of cell proliferation and protein synthesis (Lavial et al., 2007; Graf et al., 2011). In order to maintain pluripotency LIF binds to its receptor, gp130/LIFR, followed by the activation of Janus Kinase 3 (JAK) and subsequent phosphorylation of the Signal Transducer And Activator Of Transcription 3 (STAT3). Phosphorylated STAT3 interacts with Kruppel Like Factor 4 (KLF4) and maintains the pluripotency through the Octamer-Binding Transcription Factor 3/4 (OCT3/4) and its cooperatively interaction with Nanog Homeobox (NANOG) and SRY-Box Transcription Factor 2 (SOX2) (Lavial et al., 2007). In parallel to the activation of the STAT3-pathway, the binding of LIF to the LIFRβ/gp130 receptor can stimulate cell proliferation and survival by activating the MEK/ERK and PI3K/AKT signaling pathways (Figure 6 A) (Graf et al., 2011). Apart from the changes in the MEK/ERK and PI3K/AKT signaling pathways found in zfPGCs cultured for ten days, we also observed a down regulation of genes related to the maintenance of the undifferentiated state (Figure 6 B). Following up on these changes we tested the effect of LIF on zfPGC proliferation *in vitro*. We cultured zfPGCs in the presence or absence of LIF for 15 days and observed significant differences in the mean percentage of cells cultured with LIF compared to the control cultures P=0.002 N=10 (Figure 6 C).

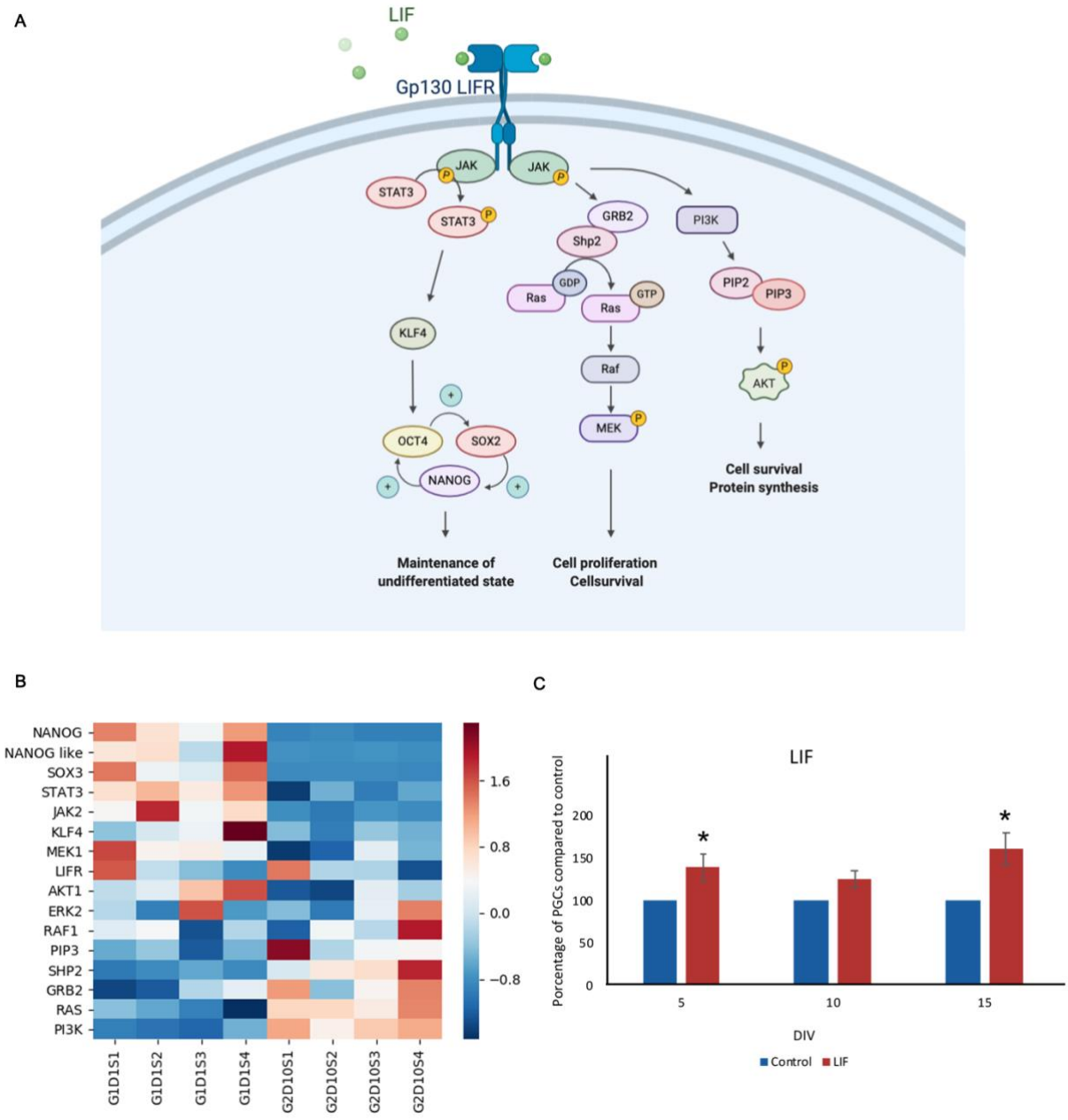


Figure 6 LIF signaling pathway and its effect on zfPGCs growth *in vitro*. (A) LIF signaling pathway. (B) Gene expression heatmaps of genes part of the LIF signaling pathway in four zfPGC samples obtained at 1 (G1S1-S4) and 10 (G2S1-S4) DIV. Z scores obtained from read counts are indicated by different colors in a range from blue (low) to red (high). (C) Effect of LIF on the growth of zfPGCs *in vitro*. Blood samples from single embryos were split in two and cultured in the presence or absence of LIF. Cell numbers were counted every 5 days and the mean percentage of cells in relation to control are given, (mean ± SEM; *P < 0.05; N=10).

TGF-ß and SCF did not show any significant effect on zfPGC proliferation in vitro

TGF-ß initiates two signaling cascades: first a canonic signaling pathway, that includes the activation of the SMAD proteins that initiate the transcription of target genes related to cell proliferation. And second non-canonical signaling cascades (non-SMAD signaling) leading to the activation of the MEK/ERK and PI3K/AKT pathways and promoting cell proliferation and survival (Weiss et al., 2012) (Figure 6 A). For its participation in cell proliferation and survival we selected this growth factor to test its effect on zfPGC proliferation *in vitro.* We cultured zfPGC in the presence and absence of TGF-ß for 15 DIV and found the differences in the percentage of cells compared to the control, and the survival time to be non-significant P=0.4 N=4 (Figure 6 B).

Stem cell factor (SCF) constitutes another important activator of both the MEK/ERK pathway as well as the PI3K/AKT pathway (Figure 1 C). When we analyzed the growth curves of zfPGCs cultured in the presence or absence of SCF for 15 DIV, we did not find significant differences P=0.4 N=7(Figure 6 C).

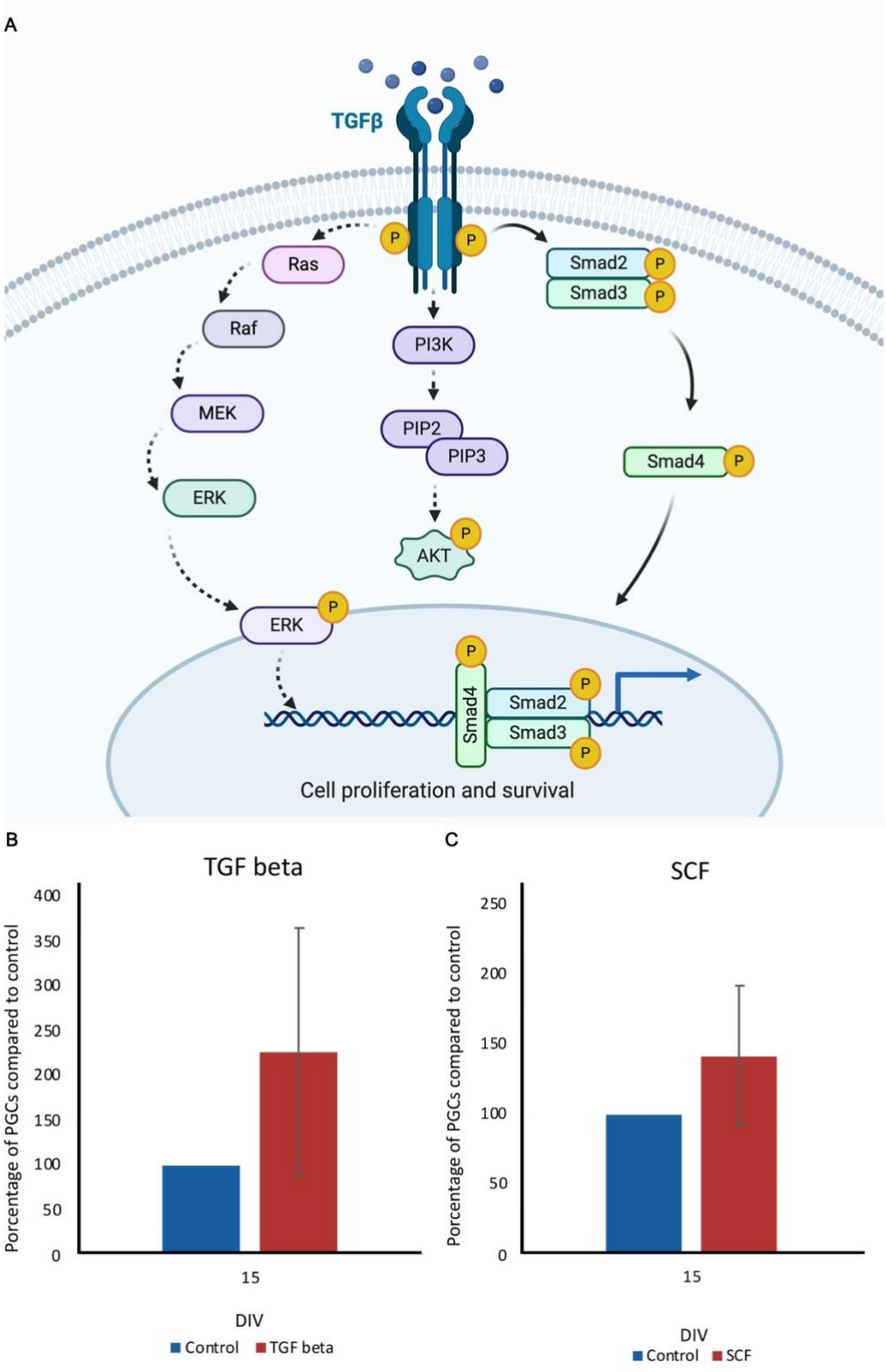


Figure 6 TGF-ß signaling pathway and zfPGC growth *in vitro* in the presence or absence of TGF-ß or SCF. (A) TGF-ß signaling pathways consisting of the canonical SMAD signaling pathway that leads to cell proliferation and the non-canonical pathways that include MEK/ERK and PI3K/AKT. (B,C) Blood samples from single embryos were

split in two and cultured in the presence or absence of TGF-ß (B) or SCF (C) . Cell numbers were counted at 1 and 15 days and the mean percentage of cells in relation to control are given (mean ± SEM; N=4 for TGF and N=7 for SCF).

Discussion

PGCs constitute a useful tool to develop transgenic avian species by their culture, genome modification and reinjection into early embryos (Van De Lavoir et al., 2006; MacDonald et al., 2010; Gessara et al., 2021). While chicken PGCs can be cultured for an unlimited time, song bird PGCs such as zfPGCs can only be cultured for a period between 2 weeks and a month (Whyte et al., 2015; Jung et al., 2019;Gessara et al., 2021). The inability to culture zfPGCs for long periods of time limits their use to produce genetic modifications by homologous recombination. ZfPGCs after being cultured for 10 days enter a stationary phase where they stop proliferating, followed by a decay phase where zfPGCs start dying. In order to determine the changes in gene expression of zfPGCs that could explain the beginning of the stationary phase we performed a comparative transcriptome analysis of zfPGCs cultured for 1 and 10 DIV.

We found about 5000 differentially expressed genes between zfPGCs cultured for 1 DIV and when cells reached the stationary phase of the growth curve at 10 DIV. We also found 20 gene ontology groups related to biological processes overrepresented in cells cultured for 1 DIV. Oxidoreductase activity acting on peroxide as acceptor was the most significant group suggesting that the ability of zfPGC of inhibiting oxidation might be affected after 10 DIV. Although hydrogen peroxide is associated with telomere shortening, G1 cell cycle arrest and apoptosis, the stationary phase does not seem to be caused by telomere shortening or cell apoptosis, since neither telomerase reverse transcriptase (TERT) nor the apoptosis pathways led by caspases show changes in the gene expression. G1 cell cycle arrest could potentially be responsible of zfPGCs entering stationary phase but our transcriptome data was inconclusive in this aspect.

The cell signaling pathways MEK/ERK, PI3K/AKT and TGF-ß are the main pathways responsible for cell proliferation and survival. These pathways were found partially up-regulated in zfPGCs cultured for 10 DIV, which could be due to the stimulus provided by the growth factors FGF2 and BMP4 present in the culture medium. However, we also

found part of these pathways to be down-regulated in zfPGCs cultured for 10 DIV. A partial down-regulation can occur in response to an attempt of the cells to maintain homeostasis in a changing environment.

Further, we identified the signaling pathways related to VEGF, EGF and LIF activity being changed in response to the culture conditions and observed a significant increase in the zfPGC number when cultured in the presence of VEGF and LIF, and a positive trend when cultured with EGF. Additionally, we tested the effects of SCF and TGF, growth factors that contribute to stimulate cell growth and proliferation, but we did not find a significant effect in the zfPGC number when comparing to the control.

LIF participates in the maintenance of an undifferentiated state of stem cells stimulating OCT 3/4 and its cooperatively interaction with NANOG and SOX2 (Lavial et al., 2007). LIF also activates the MEK/ERK and PI3K/AKT signaling pathways to stimulate stem cell proliferation and survival (Graf et al., 2011). For this reason, previous studies used LIF as part of the culture medium to successfully expand mouse and chicken PGCs *in vitro* (Resnick et al., 1992; Makoolati et al., 2016). When we tested the proliferative effect of LIF on zfPGCs in culture we found a significant increase in the cell number when compared to the control without LIF.

Apart from its participation in angiogenesis and hematopoiesis (Reviewed by Gerber et al., 2002; Breier et al., 1992; Breier et al., 1995; Flamme et al., 1995; Dumont et al., 1998), VEGF and its receptors are also present in testicular and ovarian tissue (Ergün et al., 1997; Celik-Ozenci et al., 2003) and participate in germ cell development and proliferation (Nalbandian et al., 2003, Celik-Ozenci et al., 2003, Tian et al., 2016). The proliferative effect of VEGF on zfPGFCs in culture was found in preliminary experiments not to be significant when using human VEGF (data not shown). Instead, chicken VEGF showed a significant increase in the cell number when compared to the control without VEGF. Small differences in the amino acid sequence may have been responsible for the different biological activity of human and chicken VEGF on the zfPGCs in culture when binding to the zfVEGF receptors. Comparing the amino acid sequences of human or chicken and zebra finch VEGF, we found the percentage of identity of the amino acid sequence to be 69% for human but 97.2 % for chicken VEGF when compared with zebra finch VEGF.

VEGF and LIF did not help to extend the zfPGC survival time *in vitro*, but contributed to obtain a larger number of zfPGCs in a short period of time in culture. This has two implications: first the reduction of the exposure time of zfPGCs to culture conditions and its effect on PGC homeostasis. And second the increase in the efficiency to generate transgenic songbirds allowing for a larger number of cells to be injected in zebra finch embryos to outcompete endogenous PGCs, increasing in consequence the germline transition rate.

Experimental procedures

ZfPGC collection and culture

Freshly laid zebra finch eggs were incubated for 60 hours at 37°C and relative humidity of 75% until Murray stage 13-15. Then a window was opened above the air chamber of the egg, and 1 - 3 µl of blood was extracted from the vasculature system using a pulled glass needle. Once extracted, blood samples for the 10 days group were placed in culture using the same culture conditions as in Gessara et al., 2021 for 10 days. Blood samples for the day 1 were processed after one day in culture once the samples were sexed.

Sample sexing

Only blood samples from male embryos were processed for sequencing. Following blood extraction embryonic tissue was frozen for further DNA extraction. DNA extraction was done using the DNeasy Blood and Tissue Kit (Qiagen 69504) following the manufacturer's instructions. PCR reactions were carried out with the HOT FIREPol® DNA polymerase in buffer B1 (both Solis BioDyne). A 20 µl reaction mixture included 12.5µl H_2O, 2µl 10x buffer B1, 1.6µl $MgCl_2$ (25 mM), 1µl cDNA, 2µl primer pair (10 µM), 0.4 µl dNTP mix and 0.5µl HOT FIREPol®. The thermocycling conditions were as follows: denaturation at 95°C for 15 min, 35 amplification cycles (95°C for 30 sec, 60°C for 45 sec and 72°C for 50 sec) and a final extension at 72 °C for 2 min. The amplified products were resolved by gel electrophoresis in 2% Roti®agarose (Agarose High Resolution; Roth) in 0.5x UltraPure™ TRIS-Borat-EDTA (TBE) buffer (Thermo Fisher).

The primer pair used for sexing was forward 5’ TCTGCATCGCTAAATCCTTT 3’ reverse 5’ CTCCCAAGGATGAGTAAATG 3’.

Separation of cells with Ficoll gradient

To separate zfPGCs from red blood cells a Ficoll (Sigma) gradient was prepared. Ficoll was diluted with 1ml of culture medium in two different concentrations 16 and 6.3%. Embryonic blood samples of five individuals were collected together and centrifugated at 2500 RPM. The cell pellet was resuspended in 100 ml of culture medium. 900 µl of the 16% Ficoll solution were incorporated to the cell suspension and mixed gently. In a 1.5 ml tube 200 µl of 6.3% Ficoll solution were incorporated and with a 1ml pipette the cell suspension was incorporated in the bottom of the tube underneath the 6.3% Ficoll. The tube was centrifugated at 2400 RPM for 30 minutes. With a 200 µl pipette the most superficial layer was removed (aprox 200 µl) and place it in a new 1.5 tube. A 10 µl aliquot was collected for cell counting and the rest of the cell suspension was diluted in 300 µl of culture medium. The new cell suspension was centrifugated for 5 minutes at 1800 RPM. The supernatant was carefully removed and the cell pellet was resuspended in the amount of medium necessary for RNA extraction.

RNA extraction and next generation sequencing

Four samples with 400 zfPGCs each were collected for each group (1 and 10 DIV) and RNA was extracted using the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (TAKARA) following manufactures instructions. Illumina high output sequencing was done by the NGS group at the Max Planck Institute for molecular Genetics.

Data analysis

The Galaxy web platform (https://usegalaxy.eu/) was used for the analysis of the sequencing data (Afgan et al., 2018). A quality control of the FASTQ files was performed with FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Nextera adapters were removed with Trimmomatic (Bolger et al., 2014). The reads were mapped with STAR (Dobin et al., 2013) and counted with featureCounts (Liao et al., 2014) . A quality control of the mapped sequences was done with MultiQC (Ewels et al., 2016). Differentially expressed genes between the two groups were identified with DESeq2 (Love et al., 2014).

Geneonthology analysis was done with the software Cytoscape and the Cluego application and the pathway characterization was done with the pathway database of the Kyoto encyclopedia of genes and genomes (KEGG).
The pathway diagrams were made with BioRender.com.
To generate heatmaps Z scores were calculated from read counts and heatmaps were done using the Python packages Seaborn, Pandas and Matplotlib.

Comparative culture

In order to compare the culture of zfPGCs with and without growth factors each blood sample obtained from one single embryo was divided in two, one half was cultured with the growth factor and the other half without. The growth was followed for 15 days and the cells were counted every five days. The number of zfPGCs was scored in a minimum of 8 fields under a 20× objective of an inverted microscope (Nikon eclipse Ts2). Culture of zfPGCs for comparable growth was done for Antioxidant Supplement (Sigma A1345), chicken VEGF (Biomol RP1282CT-005) 50 ng/ml, chicken LIF (Kingfisher biotech inc RP1395C-100) 50ng/ml, human EGF (Kingfisher biotech inc RP1409C) 50 ng/ml, zebra finch SCF (Kingfisher biotech inc RP1555C) 100 ng/ml and chicken TGF-ß 2 (LSBio LS-G11619-10) 50 ng/ml.

Statistical analysis

The statistical significance of the growth curves with different growth factors was done with a T-test for paired samples, all experiments were replicates at least 5 times and the level of significance was set at $P < 0.05$.

Data Records

The raw data fastq.gz files was deposited in the SRA database under the accession number PRJNA637305.

Acknowledgments

We are grateful to Anja Lohrentz, Christina Reusch, Antje Bakker, for their excellent technical support and Bernd Timmermann for performing the next generation sequencing

of our samples. IG also thanks the IMPRS research school for training and support. The study was founded by the Max Planck Society.

Supplemental information

Table 1. PI3K/AKT pathway differential expression

Pathway	Gene experession	Gene symbol	Zebra finchGene ID	Log Fold change
PI3K/AKT	Downregulated In cell with ten days in culture	ANGPT2	ENSTGUG00000012892	-2.62
		BDNF	ENSTGUG00000004743	-1.11
		CCND3	ENSTGUG00000017447	-3.60
		CCNE2	ENSTGUG00000011936	-1.57
		COL4A2	ENSTGUG00000008766	-1.30
		COL4A4	ENSTGUG00000008147	-1.25
		COL4A6	ENSTGUG00000004100	-0.94
		COL9A1	ENSTGUG00000012730	-1.47
		COMP	ENSTGUG00000003565	-1.41
		CREB1	ENSTGUG00000010023	-0.95
		CSF3	ENSTGUG00000002266	-2.06
		DDIT4	ENSTGUG00000004471	-1.01
		EFNA1	ENSTGUG00000016221	-3.71
		FGFR1	ENSTGUG00000004881	-1.93
		FN1	ENSTGUG00000002978	-1.12
		FOXO3	ENSTGUG00000012135	-1.87
		GHR	ENSTGUG00000002265	-2.54
		GNB4	ENSTGUG00000010675	-1.15
		HRAS	ENSTGUG00000009877	-1.21
		IKBKB	ENSTGUG00000005490	-1.03
		IL4R	ENSTGUG00000009509	-1.08
		ITGB5	ENSTGUG00000004062	-2.77
		JAK2	ENSTGUG00000005043	-1.68

		KIT	ENSTGUG00000007686	-3.57
		LAMB1	ENSTGUG00000003257	-1.64
		LAMB2	ENSTGUG00000008776	-1.43
		LPAR1	ENSTGUG00000001307	-2.58
		MAGI1	ENSTGUG00000009560	-1.01
		MAP2K1	ENSTGUG00000009363	-0.83
		MAP2K2	ENSTGUG00000000994	-0.92
		NR4A1	ENSTGUG00000016805	-2.38
		PDGFA	ENSTGUG00000008526	-1.26
		PDGFC	ENSTGUG00000005439	-2.39
		PPP2R3C	ENSTGUG00000015248	-1.35
		PPP2R5D	ENSTGUG00000004938	-1.45
		PRKAA2	ENSTGUG00000009569	-1.43
		PTK2	ENSTGUG00000012663	-0.70
		RBL2	ENSTGUG00000007426	-1.53
		RPS6KB2	ENSTGUG00000006119	-2.12
		RXRA	ENSTGUG00000005800	-1.71
		SGK1	ENSTGUG00000011521	-3.16
		SGK3	ENSTGUG00000011308	-1.13
		TLR2	ENSTGUG00000005179	-1.19
		TSC1	ENSTGUG00000005221	-1.21
		VEGFC	ENSTGUG00000006451	-1.63
		VTN	ENSTGUG00000005933	-2.06
		YWHAH	ENSTGUG00000009902	-1.16
	Upregulated in cells with ten days in culture	AREG	ENSTGUG00000001551	2.71
		CCNE1	ENSTGUG00000008979	1.32
		COL1A2	ENSTGUG00000001522	2.31
		COL6A1	ENSTGUG00000003341	2.03
		COL6A3	ENSTGUG00000003667	2.16
		CSF1R	ENSTGUG00000000965	2.02
		CSF3R	ENSTGUG00000001588	2.57
		CSH1	ENSTGUG00000003288	3.11

		CSH2	ENSTGUG00000003288	3.11
		EGF	ENSTGUG00000004206	1.66
		EIF4E	ENSTGUG00000003296	1.25
		EPHA2	ENSTGUG00000002048	2.02
		EREG	ENSTGUG00000001550	1.55
		FGF16	ENSTGUG00000005888	3.02
		FGF19	ENSTGUG00000005374	2.23
		FGF2	ENSTGUG00000002018	1.58
		FGF20	ENSTGUG00000007302	2.20
		FGF7	ENSTGUG00000007319	2.40
		FGF8	ENSTGUG00000009956	2.45
		FGFR4	ENSTGUG00000000438	2.81
		FLT4	ENSTGUG00000000983	1.66
		GH1	ENSTGUG00000003288	3.11
		GNB3	ENSTGUG00000013301	3.49
		HGF	ENSTGUG00000002584	1.74
		HSP90AA1	ENSTGUG00000012808	1.11
		IL6	ENSTGUG00000002725	2.44
		IL7R	ENSTGUG00000001996	2.61
		ITGA1	ENSTGUG00000002382	1.33
		ITGB4	ENSTGUG00000008343	1.49
		ITGB6	ENSTGUG00000002009	2.22
		KRAS	ENSTGUG00000012235	0.86
		LAMA4	ENSTGUG00000011973	1.46
		LAMC3	ENSTGUG00000004775	2.04
		MTOR	ENSTGUG00000002263	1.20
		PCK1	ENSTGUG00000008254	3.00
		PDGFRB	ENSTGUG00000000942	2.17
		PIK3CA	ENSTGUG00000010736	2.02
		PPP2CB	ENSTGUG00000001762	0.79
		PRLR	ENSTGUG00000001979	1.99
		RAS	ENSTGUG00000012235	0.86

		RPTOR	ENSTGUG00000003028	0.82
		TGFA	ENSTGUG00000015405	1.48
		THBS1	ENSTGUG00000011624	2.67
		THBS4	ENSTGUG00000014102	1.65
		TLR4	ENSTGUG00000003342	3.25
		TNR	ENSTGUG00000017420	2.51
		VEGFA	ENSTGUG00000008137	1.75
		YWHAZ	ENSTGUG00000012078	1.32

Table 2. MEK/ERK pathway differential expression

Pathway	Experession profile	Gene Symbol	Zebra finchGene ID	Log Fold change
MEK/ERK	Downregulated in cells with ten days in culture	ANGPT2	ENSTGUG00000012892	-2.62
		BDNF	ENSTGUG00000004743	-1.11
		BRAP	ENSTGUG00000006520	-0.96
		CALML4	ENSTGUG00000009596	-1.34
		CDC42	ENSTGUG00000017253	-1.09
		EFNA1	ENSTGUG00000016221	-3.71
		FGFR1	ENSTGUG00000004881	-1.93
		GAB1	ENSTGUG00000002400	-0.98
		GAB2	ENSTGUG00000013005	-1.85
		GNB4	ENSTGUG00000010675	-1.15
		GNG12	ENSTGUG00000010203	-1.56
		GNG4	ENSTGUG00000010286	-2.81
		HRAS	ENSTGUG00000009877	-1.21
		IKBKB	ENSTGUG00000005490	-1.03
		JMJD7-PLA2G4B	ENSTGUG00000011002	-1.90

		KIT	ENSTGUG00000007686	-3.57
		MAP2K1	ENSTGUG00000009363	-0.83
		MAP2K2	ENSTGUG00000000994	-0.92
		MRAS	ENSTGUG00000004727	-1.04
		PAK1 p1 (RAC1)	ENSTGUG00000013034	-1.18
		PDGFA	ENSTGUG00000008526	-1.26
		PDGFC	ENSTGUG00000005439	-2.39
		PLA2G12B	ENSTGUG00000004413	-2.32
		PLA2G4B	ENSTGUG00000011002	-1.90
		RAC2	ENSTGUG00000010650	-2.12
		RAC3	ENSTGUG00000003640	-2.00
		RALGDS	ENSTGUG00000005351	-1.41
		RAPGEF5	ENSTGUG00000002700	-1.56
		REL	ENSTGUG00000006743	-1.87
		SHC1 SHC	ENSTGUG00000004152	-0.84
		VEGFC	ENSTGUG00000006451	-1.63
	Upregulated in cells with 10 days in culture	ARF6	ENSTGUG00000013196	0.93
		CALM1	ENSTGUG00000012465	1.78
		CALM2	ENSTGUG00000002581	0.97
		CSF1R	ENSTGUG00000000965	2.02
		EGF	ENSTGUG00000004206	1.66
		EPHA2	ENSTGUG00000002048	2.02
		FGF16	ENSTGUG00000005888	3.02
		FGF19	ENSTGUG00000005374	2.23
		FGF2	ENSTGUG00000002018	1.58
		FGF20	ENSTGUG00000007302	2.20

		FGF7	ENSTGUG00000007319	2.40
		FGF8	ENSTGUG00000009956	2.45
		FGFR4	ENSTGUG00000000438	2.81
		FLT4	ENSTGUG00000000983	1.66
		GNB3	ENSTGUG00000013301	3.49
		GRIN1	ENSTGUG00000002568	2.32
		GRIN2A	ENSTGUG00000004747	1.77
		GRIN2B	ENSTGUG00000009456	1.84
		HGF	ENSTGUG00000002584	1.74
		JMJD7-PLA2G4B	ENSTGUG00000011000	2.49
		KRAS	ENSTGUG00000012235	0.86
		MAPK10	ENSTGUG00000002750	1.66
		PAK2; p21 (RAC1)	ENSTGUG00000005832	1.39
		PDGFRB	ENSTGUG00000000942	2.17
		PIK3CA	ENSTGUG00000010736	2.02
		PLA1A	ENSTGUG00000013384	2.24
		PLA2G12A	ENSTGUG00000004164	1.52
		PLA2G1B	ENSTGUG00000009518	2.38
		PLA2G4B	ENSTGUG00000011000	2.49
		PLA2G4D	ENSTGUG00000011006	1.71
		PLA2G4F	ENSTGUG00000011023	2.32
		PLCG2	ENSTGUG00000004405	1.51
		PTPN11	ENSTGUG00000006795	2.05
		RAB5A	ENSTGUG00000003236	0.81
		RALBP1	ENSTGUG00000009818	1.38
		RASA3	ENSTGUG00000009356	1.25

		RRAS2	ENSTGUG00000008422	1.10
		SHC3	ENSTGUG00000000448	1.48
		SHC4	ENSTGUG00000007228	1.09
		SYNGAP1	ENSTGUG00000017277	2.52
		TGFA	ENSTGUG00000015405	1.48
		TIAM1	ENSTGUG00000013461	1.27
		VEGFA	ENSTGUG00000008137	1.75
		ZAP70	ENSTGUG00000000330	1.44

Table 3. SMAD pathway differential expression

Pathway	Experession profile	Gene Symbol	Zebra finchGene ID	Log Fold change
SMAD	Downregulated in cells cultured for 10 days	BMP7	ENSTGUG00000008348	-2.67
		GREM2	ENSTGUG00000010595	-2.01
		ID2	ENSTGUG00000012998	-1.70
		ID4	ENSTGUG00000006296	-1.75
		LTBP1	ENSTGUG00000008923	-1.10
		NEO1	ENSTGUG00000004132	-0.78
		RGMA	ENSTGUG00000014262	-1.62
		RPS6KB2	ENSTGUG00000006119	-2.12
		SMAD1	ENSTGUG00000002675	-1.00
		TGFB2	ENSTGUG00000002825	-0.88
		SMAD8	ENSTGUG00000011897	-1.36
	Upregulated in cells cultured for ten days	ACVR1C	ENSTGUG00000016862	2.15
		CHRD	ENSTGUG00000010366	2.42
		E2F5	ENSTGUG00000011713	1.67
		GDF5	ENSTGUG00000003594	1.96
		LEFTY1	ENSTGUG00000004196	2.55
		PPP2CB	ENSTGUG00000001762	0.79
		ROCK1	ENSTGUG00000010253	1.28

		SKP1	ENSTGUG00000001291	0.99
		SMAD2	ENSTGUG00000000016	0.75
		SMAD6	ENSTGUG00000009467	0.97
		THBS1	ENSTGUG00000011624	2.67
		SMAD3	ENSTGUG00000009470	1.18

Chapter III

Determination of expression and functionality of two novel zebra finch androgen receptor splice variants

Ivana Gessara

Falk Dittrich

Moritz Hertel

Manfred Gahr

Manuscript in preparation for publication

Expression and activity of two novel zebra finch androgen receptor splice variants

Authors

Ivana Gessara[1]*, Falk Dittrich[1], Moritz Hertel[1], Manfred Gahr [1]

1. Max Planck Institute for Ornithology, Eberhard Gwinner Strasse, 82319 Seewiesen, Germany

Abstract

The androgen receptor (AR) plays an essential role in the development of the song control nuclei in the brain of songbirds. Few studies had looked into the structure of the zebra finch AR (zfAR) gene and AR splice variants have not been reported yet. Here, I have investigated the expression pattern of the two novel AR splice variants zfAR-V2-4 and zfAR-V2-8 by RT-PCR and assessed their activity in an *in vitro* AR reporter assay. *In vitro,* zfAR-V2-4 showed a ligand-independent genomic effect of 5% when compared to the full length zfAR (zfAR-FL), while zfAR-V2-8 had an inhibitory effect of 55% on the genomic activity of the zfAR-FL similar to the synthetic human AR inhibitor variant ARi410 and its zebra finch homologue. Together with the method developed by Gessara et al. (2021) the availability of an inhibitory AR variant opens the possibility to develop a transgenic dominant-negative knock-down model in the zebra finch.

Introduction

The AR plays an essential role in brain development in mammals and birds. Studies in rodents showed that testosterone in particular, through its interaction with the AR, stimulates cell survival (Spritzer and Galea, 2007; Hamson et al., 2013) and synaptogenesis in both males and females (Leranth et al., 2003). It also regulates the volume of the amygdala (Cooke et al., 1999), soma size and dendritic length of adult motor neurons (Kurz et al., 1986), and the density of dendritic spines in the male preoptic area, a brain area that is involved in the control of sexual behaviours (Garelick and Swann, 2014).

The AR is a sex hormone receptor that is activated after binding of testosterone or 5α-dihydrotestosterone (DHT) (Grino et al., 1990) and can show two different mechanisms of activity including canonical and non-canonical signalling pathways. In the absence of ligands, the AR is located in the cytoplasm where binding of testosterone or DHT elicits conformational changes so that the AR-androgen complex is translocated into the cell nucleus. There the AR dimerizes and binds onto the DNA recognizing androgen response elements (AREs) present in the promoter region of AR target genes to modulate the expression (Edler et al., 2001). Co-regulators can bind to the activated AR and enhance or repress its genomic effects (Bevan et al., 1999). This process constitutes the canonical

signalling of the AR. The non-canonical signalling of the AR is based on DNA binding independent mechanisms by activating second messenger signalling pathways that include the Extracellular-signal Regulated Kinases (ERK), Serine/Threonine kinases (AKT) and Mitogen-Activated Protein Kinases (MAPK) (Estrada et al., 2003, Kang et al., 2004, Kousteni et al., 2001, Gill et al., 2004).

The gene structure of the AR is the most widely studied in humans where it consists of eight protein-coding exons. In the full-length human AR (hAR-FL) exon 1 encodes the amino terminal domain (NTD) (Ferro et al., 2002; Ding et al., 2004 and 2005), exon 2 and 3 the DNA binding domain (DBD) (Shaffer et al., 2004), exon 4 the hinge region (HR) and exons 5 to 8 the ligand binding domain (LBD) (Matias et al., 2000). It has two transactivation regions, one ligand independent, AF-1, located in the NTD (Callewaertand et al., 2006) and one ligand dependent, AF-2, located in the LBD (Slagsvold et al., 2000). Furthermore AF-1 contains two transcription activation units Tau-1 and Tau-5. Interestingly, Tau-5 retains the activation potential in the absence of the LBD while Tau-1 depends on it to be activated (Jenster et al., 1995). The AR also contains a nuclear localization signal (NLS) that imports the receptor into the nucleus and a nuclear export signal (NES) responsible for moving the receptor back to the cytoplasm in case of ligand withdrawal (Tan et al., 2015) (Figure 1).

More than 30 distinct AR splice variants have been identified so far (Lu et al., 2020; Wach et al., 2020) mainly in human prostatic cancer cell lines (Ahrens-Fath et al., 2005; Haile et al., 2011; Lu et al., 2013; Cao et al., 2016). The naturally occurring splice variant hAR45 (Ahren-Fath et al., 2005; Weiss et al., 2007; Hu et al., 2014; Garza-Conteras et al., 2017) was first identified to have an inhibitory effect on the hAR-FL activity *in vitro* after heterodimerization (Ahren-Fath et al., 2005) and recently speculated to play a role as a membrane androgen receptor that acts via fast non-genomic signalling (Ahren-Fath et al., 2005; Schreihofer et al., 2018; Duong et al., 2020). In comparison to the hAR-FL, in hAR45 exon 1 is replaced by a short unique N-terminal extension of seven amino acids, the exon 1B, using an alternative promoter downstream of exon 1 (Figure 1) which is encoded by a sequence between exon 1 and 2. Additional natural occurring variants like hAR-V1, hAR-V3, hAR-V7 and hAR-V9 (Hu et al., 2009; Watson et al., 2010) enclose N-terminally the NTD with the transcription activation units followed by the DBD but lack both the HR and C-terminally the LBD resulting in constitutive (hAR-V3 and -7) or conditional activity (Wach et al., 2020) (Figure1). Their C-terminal ends are encoded by

cryptic exons (CE) that arise from non-canonical splicing events and code for a relative short amino acid sequence.

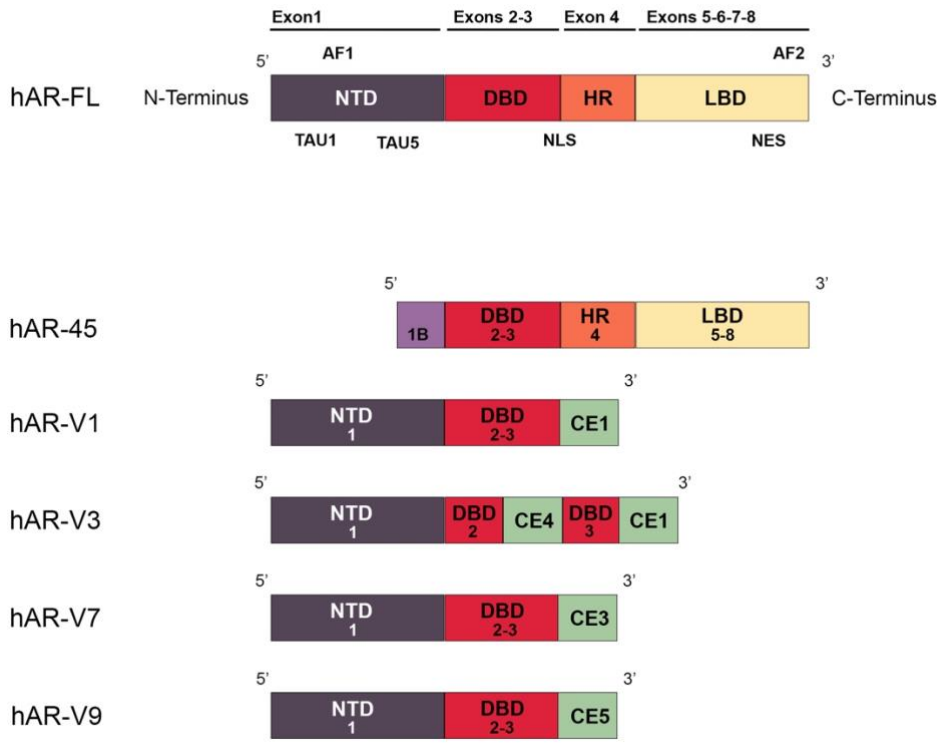
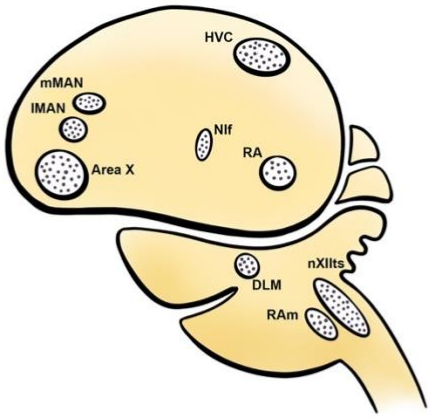


Figure 1. Human androgen receptor transcripts. In the human full-length androgen receptor (hAR-FL) the amino terminal domain (NTD) is encoded by exon 1 and contains the Tau1 and Tau5 transactivation sequences within the transactivation region AF-1. A DNA binding site (DBD) is encoded by exons 2 and 3, the hinge region (HR) by exon 4 and the ligand binding domain (LBD) by exons 5 to 8 the latter containing the transactivation region AF-2. A nuclear localization signal (NLS) and a nuclear export signal (NES) can be found between the DBD and the HR and within the LBD, respectively. The hAR45 constitutes the most highly expressed natural occurring splice variant of the hAR. It differs from the hAR-FL transcript in that exon 1 is replaced by 1B. The hAR splice variants hARV 1, 3, 7 and 9 are less frequently expressed, lack both HR and LBD, and contain the cryptic exons CE) 1, 3, 4 or 5.

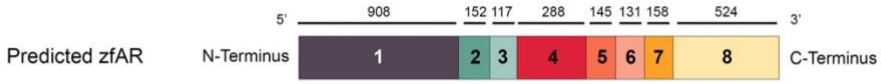
In songbirds, the AR plays important roles in the development (Gahr, 2004) and adult plasticity (Chen et al., 2013) of the song control system as well as in song learning, production and perception (Cynx and Nottebohm, 1992; Cynx et al., 2005; Apfelbeck et al., 2012). In song control system of zebra finches the AR was found to be expressed in HVC, mMAN, lMAN, Area X, Nif, RA, DLM, RAm and nXIIts (Figure 2 A) (Balthazart et al., 1992; Bernard et al., 1999; Gahr et al., 1998, 2008; Metzdorf et al., 1999; Fusani et al., 2000; Voigt and Gahr, 2011). The structure of the zebra finch AR gene was predicted by GENSCAN analysis of a BAC library made with female zebra finch livers by Luo et al. in 2006. The predicted full-length zebra finch AR (zfAR-FL) transcript was expected to be formed by 2423 base pairs that encode eight axons (Figure 2 B). However, when performing a 5'-RACE experiment Luo et al. (2006) could not amplify the first 619 base pairs of exon 1 (Figure 2 C) (Luo et al., 2006). Recently, two promoter regions in front of exon 1 and 2 of the zfAR gene were detected by bioinformatic analysis (Frankl-Vilches et al., 2017) and in experiments using the Rapid amplification of cDNA ends (RACE) technique (Figure 2 D; Dittrich and Lohrentz, unpublished results). Furthermore, two novel zfAR variants, zfAR-V2-4 and zfAR-V2-8, could be identified (Dittrich and Lohrentz, unpublished results) that use the alternative promoter located upstream of exon 2. Both lack the exon 1 but contain a unique N-terminus of 26 amino acids encoded by the immediate 5'untranslated region (UTR) of exon 2 called exon 1b (Dittrich and Lohrentz, unpublished results; Figure 3 b, C). The zfAR-V2-8 apart from exon 1b also contains exons 2 to 8 including the DBD, HR and LBD, while the zfAR-V2-4 contains exons 1b, 2, 3, 4 and a unique C-terminus of 59 amino acids but lacks the LBD (Dittrich and Lohrentz, unpublished results; Figure 3).

In this study I have investigated the expression of the novel zfAR splice variants zfAR-V2-4 and zfAR-V2-8 by RT-PCR in different zebra finch tissues *in vivo* and their functionality *in vitro* in an AR reporter assay. I found the splice variants zfAR-V2-4 and zfAR-V2-8 to be widely expressed *in vivo* and to exhibit *in vitro* a low transactivational effect (zfAR-V2-4) and a substantial inhibitory effect (zfAR-V2-8), respectively. For comparison, the inhibitory effect of the synthetic human dominant-negative AR variant ARi410 and its zebra finch homologous (zfARi) on the activity of the zfAR-FL in the AR reporter assay was tested as well.

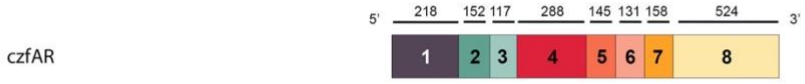
A



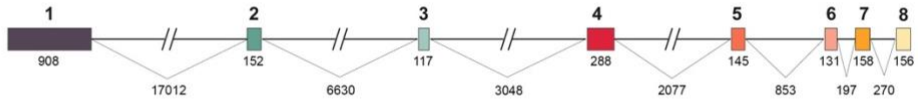
B



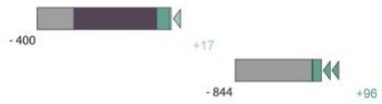
C



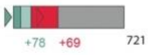
D



5' RACE



3' RACE



E

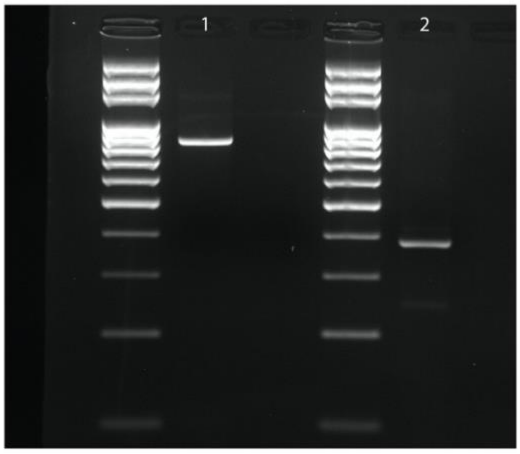


Figure 2. Expression in song control nuclei and alternative transcripts of the zfAR gene. (A) Representation of the songbird brain showing nuclei of the song control system that contain cells expressing the androgen receptor (dots) in HVC, mMAN, lMAN, Area X, Nif, RA, DLM, RAm and nXIIts. (B) Predicted zfAR-FL transcript with 1 to 8 (Perlman et al., 2003). (C) CzfAR transcript reported for the liver of female zebra finches when detected by 5'-RACE analysis (Luo et al., 2006). From exon 1 only the last 218 base pairs could be amplified. In B, C numbers above exons indicate exonal length but exon 8 includes the coding sequence and the 3' untranslated region (UTR). (D) Gene structure of the zfAR with exons 1 to 8; length of exons and introns are indicated. Products obtained by 5' and 3' RACE analyses are shown and gene specific primers that were used are indicated by arrow heads. Length of UTRs are given. Note, that the zfAR gene includes two different promoter regions upstream of exon 1 and 2, respectively. (Figure D is from Dittrich and Lohrentz, unpublished results). (E) *Picture of PCR-Gel with amplicons for* (1) *Exon 1, 5' UTR exon 2 and* (2) *3'UTR 4.*

Results

Tissue expression of the zfAR splice variants zfAR-V2-4 and zfAR-V2-8.

RT-PCRs were performed with primers that bound to sequences located in the 5'-UTR of exon 1b for the forward primers and in 3'-UTRs for the reverse primers that were downstream of exon 4 in the case of zfAR-V2-4 and exon 8 in the case of zfAR-V2-8, respectively (Figure 3 A-C). Expression of the two novel zfAR variants zfAR-V2-4 and zfAR-V2-8 could be detected in several different tissues and in many brain tissues at different developmental stages similarly to the zfAR-FL (Table 1; Figure 3 D). However, unlike zfAR-V2-8, the AR-V-2-4 was not found to be expressed in kidney, muscle and heart. All variants were found to be most strongly expressed in testis and for zfAR-FL and zfAR-V2-4 the expression level seems to increase in the brain from post hatch day one to adulthood (Table 1; Figure 3 D).

Table 1. Expression of AR variants in zebra finch tissues.

Tissue	AR-V-2-4	AR-V-2-8
Forebrain male 1 DPH	+	++
Forebrain female 1 DPH	+	++
Forebrain male 10 DPH	++	++
Forebrain female 10 DPH	++	+
Forebrain adult male	++	+++
HVC 30 DPH	+++	++
HVC 50 DPH	++	++
HVC 100 DPH	++	++
Testis	+++	+++
Liver	++	++
Kidney	-	+
Lung	++	++
Heart	-	+
Muscle	-	+

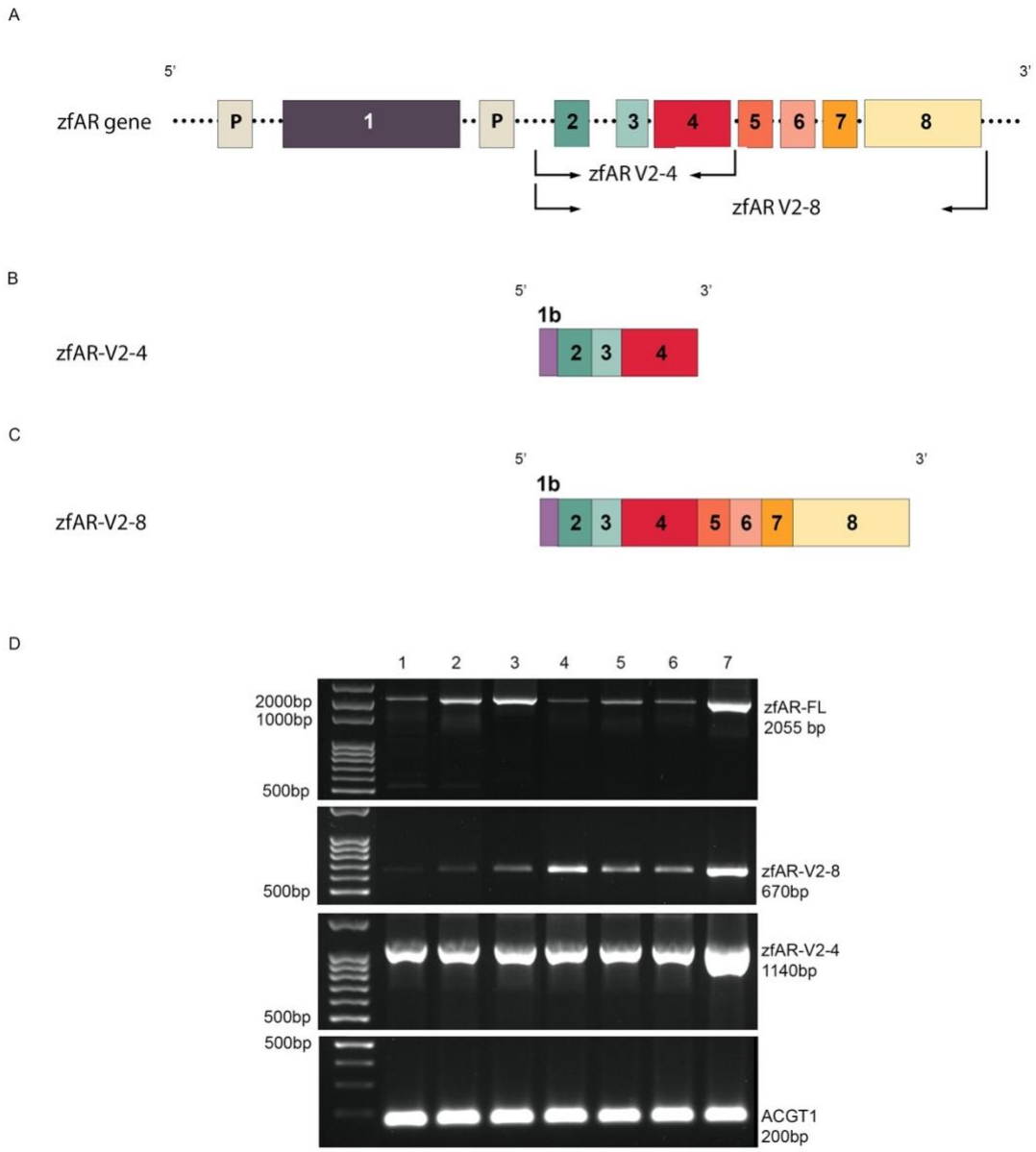


Figure 3. Detection of the two novel zfAR variants zfAR-V2-4 and zfAR-V2-8 by RT-PCR (A) In the zfAR gene promoter regions (P) are present upstream of exon 1 and exon 2, (Frankl-Vilches and Gahr, 2017; Dittrich and Lohrentz, unpublished results). Primers used for the detection of the two novel zfAR variants by RT-PCR were specific for the common 5'-UTR of exon 2 and the transcript specific 3'-UTR sequences. (B, C) Exonal structure of the two novel zfAR variants. (D) Photographs of agarose gels after electrophoretic separation and ethidium bromide staining of RT-PCR products. ZfAR full length and variants 2-4 and 2-8 (zfAR-FL, zfAR-V2-4 and zfAR-V2-8) and y-actin (ACGT1 positive control) (ACTG1) were detected in the forebrain at (1) 1day post

hatching (DPH), (2) 10 DPH and (3) of adults, the HVC (4) at 30 DPH, (5) 50 DPH and (6) 100 DPH and (7) testis.

Constitutive and ligand-dependent transactivational activity of the zfAR-V2-4

To determine the functionality of the novel zfAR splice variants zfAR-V2-4 and zfAR-V2-8 luciferase AR promoter assays were performed *in vitro* with the HeLa cell line measuring the transactivational activity in the presence or absence of the native ligand 5-dihydrotestosterone (DHT). The zfAR-V2-8 did not induce any luciferase activity neither in the absence nor in the presence of DHT. In contrast, the zfAR-V2-4 elicited alone 5 % and in the presence of DHT 27 % of the transactivational activity of the zfASR-FL (Figure 4 A). Since HeLa cells do not express the AR (Tararova et al., 2007), heterodimerization of the zfAR-V2-4 and endogenous AR encoded by the HeLa cells cannot be responsible for the genetic activity. Thus, the zfAR-V2-4 variant was constitutively active to a certain extent and showed considerable transactivational activity in the presence of the ligand. This finding was surprising, because the zfAR-V2-4 lacks exon 1 which harbours the transactivation units 1 and 5 in the hAR, and, in particular, lacks the ligand binding domain that is encoded by exons 5 to 8.

Inhibitory zfAR variants

To evaluate transgenes that could be used for the generation of a transgenic zebra finch model that expresses a dominant-negative AR variant the effect of the previously developed human synthetic dominant-negative variant ARi410 (Butler et al., 2006) and its zebra finch homologue (zfARi) on the transactivational activity of the zfAR-FL were tested in the same AR luciferase reporter assay. Both variants lack most of exon 1 but include exons 2 to 8. ARi410 was originally developed as a therapeutic approach to treat prostatic cancer. Butler and colleagues discovered that a variant of the human AR lacking the transactivation sequences AF-1 and AF-5 from the NTD in exon 1 will compete with the AR-FL for binding to the ligand, but will not produce a transactivational effects, reducing in consequence the genomic activity of the AR-FL. When testing the zfAR-FL together with the ARi410 the induction of the luciferase activity by DHT was reduced by

up to 72 %. Similarly, with the synthetic zfARi a reduction of up to 67 % was obtained (Figure 4 B).

Since the native zfAR splice variant zfAR-V2-8 lacks exon 1 as well its inhibitory effect was investigated next. Co-transfection of HeLA cells with the zfAR-FL and zfAR-V2-8 resulted in a reduction of the luciferase activity in the presence of DHT by 55% indicating that the zfAR-V2-8 can act as an inhibitory zfAR splice variant (Figure 4 B). However the inhibitory effect was stronger by more than 10 % when using the synthetic variants. In contrast and as expected from the former results an increase of the transactivational effect by approximately 30% could be observed when the zfAR-FL and the zfAR-V2-4 were co-transfected indicating a synergistic effect (Figure 4 B).

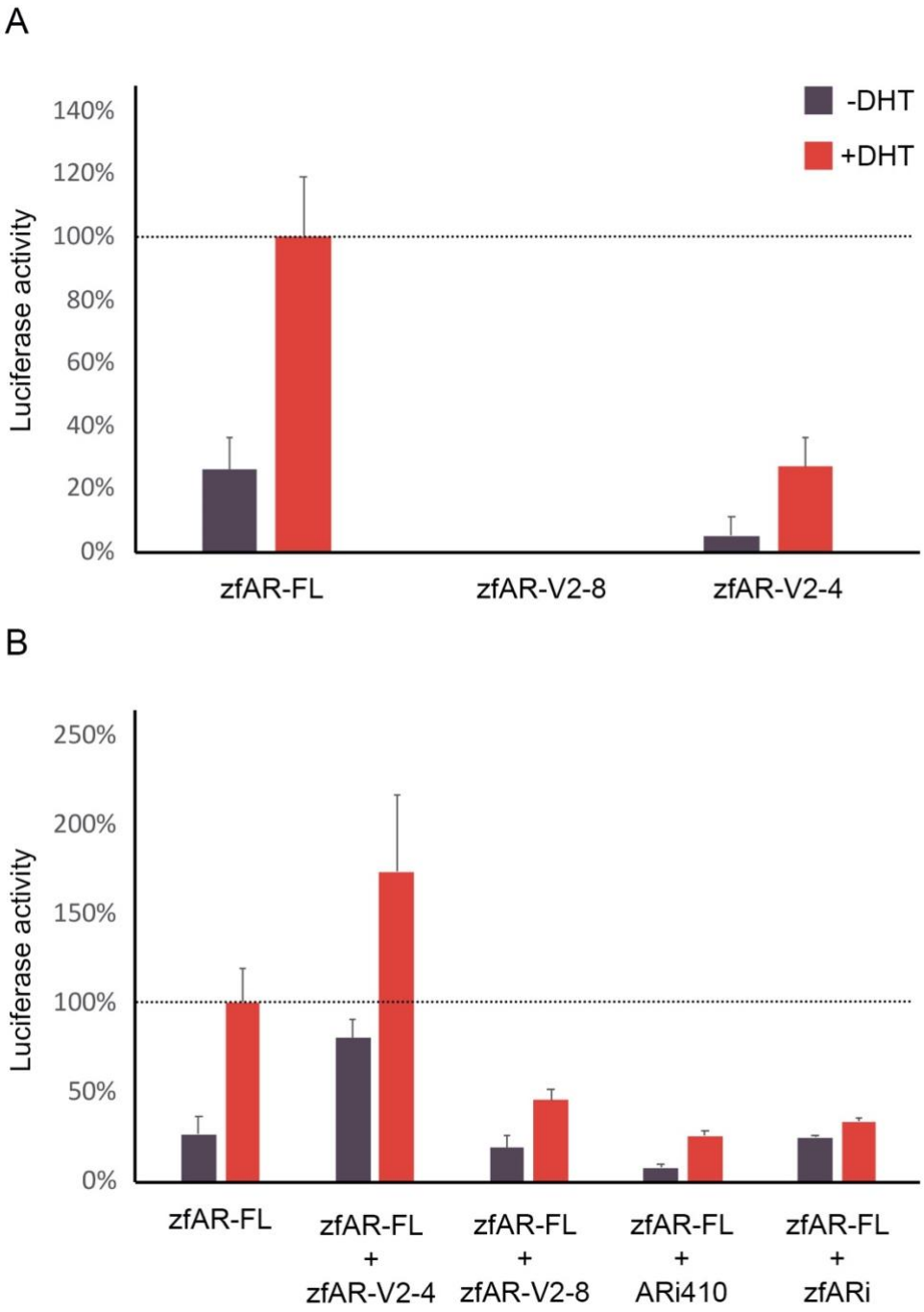


Figure 4. Transactivational activity of zfAR variants in an luciferase AR reporter assay. HeLa cells were transfected with an luciferase AR reporter plasmid and expression plasmids for different zfAR variants. Resulting luciferase activity was determined in the absence (purple bars) and presence (red bars) of 1nM 5α-dihydrotestosterone (DHT). Data are presented as percentage of the luciferase activity induced by zfAR-FL (mean ± SEM of six cultures). (A) HeLa cells were transfected with zfAR-FL or the native splice variants zfAR-V2-4 and zfAR-V2-8. (B) Co-transfection of zfAR-FL together with the

synthetic human variant ARi410, its zebra finch homologue ARi, or one of the two novel native splice variants of the zfAR.

Discussion

The androgen receptor constitutes a key regulator of many developmental and behavioural processes including the development of the song control system in songbirds (Gahr, 2004; Chen et al., 2013). In relation to the gene structure and function of the AR the zebra finch is one of the most studied songbird species. In 2005, Luo and colleagues (Luo et al., 2005) predicted the zfAR to enclose 8 axons and identified by RACE experiments the variant czfAR that starts at 690 bp of exon 1 and extends to exon 8. In 2017 bioinformatic analyses of Frankl-Vilches and Gahr revealed two promoter regions in the zfAR gene, one in front of exon 1 and the other in front of exon 2. Recently, Dittrich and Lohrentz (unpublished results) detected by RACE and RT-PCR two zfAR splice variants that seems to be expressed under control of a cryptic promoter in front of exon 2: zfAR-V2-4 and zfAR-V2-8. In this study, I have assessed the expression of these two novel zfAR splice variants in different tissues and found them to be expressed in testis, liver, lung and brain particularly in the male song control nucleus HVC during the song learning phase and adulthood. However, unlike zfAR-V2-8, expression of zfAR-V2-4 was not found in kidney, heart and muscle and its expression level seems to increase in the brain from day one post hatching to adulthood.

In order to test the functionality of the zfAR splice variants they were monitored *in vitro* using a luciferase AR reporter assay. When testing the zfAR-V2-4 both ligand-independent and -stimulated transactivational activity could be observe. Thus, the zfAR-V2-4 can be constitutively active generating a small effect in the absence of the ligand and shows enhanced activity in the presence of the ligand. Such ligand-independent activity was previously reported for human AR splice variants such as AR-V1, 3, 7 and 9 that, like zfAR-V2-4, lack the LBD and are also found to be expressed in normal tissues (Gregory et al., 2001; Ceraline et al., 2003; Dehm et al., 2008; Sun et al., 2010; Haile et al., 2011; Lu et al., 2013; Hu et al., 2014). However, in contrast to these human variants the zfAR-V2-4 lacks exon 1 that was shown to harbour an activation function domain. Instead, co-factors might be responsible for the transactivational effects of the zfAR-V2-4. Such cofactor-dependent activity *in vitro* is known from the human splice variant hAR-

45 that lacks exon 1 as well and shows activity in a promoter assay in the presence of ß-catenin or the nuclear receptor coactivator 2 (Ahren-Fath et a., 2005).

Although the zfAR-V2-4 also lacks the LBD it showed a ligand-dependent increase of the activity. Since HeLa cells do not express the AR (Tararova et al., 2007), heterodimerization of the zfAR-V2-4 and endogenous AR encoded by the HeLa cells cannot be responsible for the enhanced activity. Instead, a new LBD might be present in the 3' UTR of exon 4. Future studies of a zfAR-V2-4 with deletions and mutations in the 3'-UTR of exon 4 could help to investigate this possibility. Interestingly, when both zfAR-V2-4 and zfAR-FL were included in the same promoter assay a synergistic effect on the luciferase activity occurred, indicating a potential heterodimerization of the zfAR-V2-4 and the zfAR-FL. The heterodimerization of AR variants with the AR-FL was already reported in humans, where AR variants such as AR-V7 that also lack the LBD bind with the AR-FL by N- and C-terminal interactions resulting in the transactivation of target genes (Xu et al., 2015). Since the zfAR-V2-4 that was used here contained a nine amino acid hemagglutinin (HA) tag for the purpose of immunohistological detection when applied in vivo additional tests need to be done to exclude transactivational and ligand binding functions of the HA-tag.

When testing the splice variant zfAR-V2-8 together with the zfAR-FL in the promoter assay, a 55 % reduction of the transactivational effect of the zfAR-FL was found. Indicating that zfAR-V2-8 is an inhibitory zfAR splice variant that is ubiquitously expressed. Strikingly, another natural occurring AR variant was discovered in humans (hAR-45) to have an inhibitory effect on the hAR-FL activity (Ahren-Fath et al., 2005). Structurally, both hAR-45 and zfAR-V2-8 lack exon 1 that is usually responsible for transactivational effects, but use a promoter region located between exon 1 and 2 to express a short sequence that constitutes exons 1B and 1b, respectively. Since the zfAR-V2-8 did not exhibited any constitutive activity it may be used as a naturally occurring candidate for the generation of a dominant-negative transgenic zebra finch model to diminish the effect of androgens during the development of the song control system. Additionally, the synthetic inhibitory AR variants ARi410 and its zebra finch homologue were tested in this study on reduction of the transactivational effect of the zfAR-FL. Their inhibitory effect turned out to be larger by more than 10 % compared to the effect of the zfAR-V2-8. In conclusion, both naturally occurring zfAR-V2-8 and the synthetic

inhibitory AR variants constitute a viable tool for the generation of a transgenic zebra finch with an AR knock-down.

To assess potential effects of the new zfAR variants on the activity of the zfAR-FL in different tissues and at different developmental stages the relative copy numbers of the transcripts and protein concentrations need to be determined. In human tissues the hAR45 was reported to be generally the most abundant AR variant (Hu et al. 2014), strongly expressed in the heart (Ahrents-Fath et al., 2005) where its mRNA expression level approaches about 20 % of the hAR-FL and to reach about 4 % in the brain and 5 % in the placenta (Hu et al. 2014). Despite of the relative low expression level compared to the hAR-FL in human placenta increased expression of hAR45 in asthmatic patients was reported to be associated with an enhanced expression of androgen regulated target genes including the vascular endothelial growth factor (VEGF) (Meakin et al., 2019). This growth factor plays an important role in the testosterone-dependent differentiation of the song control nuclei in canaries (Louissaint et ., 2002; Hartog et al., 2009). Thus, other songbird species like the canary, that show significant androgenic effects on the differentiation of the song control system and singing behaviour should be considered as a useful model to study the biological roles of the new songbird AR variants.

Experimental procedures

RNA extraction

Total RNA was isolated from different male tissues like testis, liver, lung, heart and skeletal muscle. Additionally from the forebrain at embryonic day 8, post hatching days 1 and 10, as well as from adult birds. Male HVC was cut out under a stereo microscope from cryostat section of forebrains that were obtained at 30, 50 and 100 days after hatching as well as in adulthood. For total RNA isolation the RNeasy mini kit (Qiagen) was used following the manufacturer’s instructions.

cDNA synthesis

For cDNA synthesis, 2 - 3 µg of total RNA were denaturised in 10 µl distilled water in the presence of 1µl random hexamers (50 ng/µl) and 1µl desoxyribonukleosidtriphosphates (dNTP mix; 10 mM each) for 5 minutes at 65°C. After chilling the solution on ice for 1 min, I added 2µl 10x reverse-transcriptase (RT) buffer, 4µl 25 mM MgCl2, 2µl 0.1 M 1,4-dithio-D-threitol (DTT), 1µl RNaseOUTTM (40U/µl; Thermo Fisher) and 1µl of the RT SuperScriptTM III (200U/µl; Thermo Fisher). Reverse transcription was carried out in a thermocycler for 10 min at 25°C, followed by 50 min at 50°C and, finally, 5 min at 85°C. Next, the reaction mix was placed on ice for 5 min and, after 1µl RNase H (2U/µl; Thermo Fisher) was added, incubated for 20 min at 37°C.

RT-PCR analyses

PCR reactions were carried out with the HOT FIREPol® DNA polymerase in buffer B1 (both Solis BioDyne). A 20 µl reaction mixture included 10.5µl H2O, 2µl solution S (Solis BioDyne), 2µl 10x buffer B1, 1.6µl MgCl2 (25 mM), 1µl cDNA, 2µl primer pair (10 µM), 0.4 µl dNTP mix and 0.5µl HOT FIREPol®. The thermocycling conditions were as follows: denaturation at 95°C for 10 min, 35 amplification cycles (95°C for 30 sec, 65°C for 40 sec and 72°C for 30 sec) and a final extension at 72 °C for 5 min. The amplified products were resolved by gel electrophoresis in 2% (Roti® Agarose High Resolution; Roth) in 0.5x UltraPureTM TRIS-Borat-EDTA (TBE) buffer (Thermo Fisher). The primer pair designed for zfAR-FL was forward 5'-CGGGGCTTTTCAGAGCTTCTT -3' and reverse 5'-AAGTAACAGGCAGGAGATGGC -3', for zfAR-V2-4 was forward 5'-TGCAGCAGCCAGTCCCAGTCGC-3' and reverse 5'-GAGTCCAGACCAGCTGCAGCCCA-3'and for zfAR-V2-8 was forward 5'-GATCTGTGGGGATGAAGC-3' and reverse 5'-AAGTAACAGGCAGGAGATGGC-3'.

Luciferase AR reporter assay

HeLa cells were maintained *in vitro* in DMEM high glucose (Gibco 31966-021) containing 10 % fetal calf serum (FCS). For each experiment $1.5x10^4$ HeLa cells were seeded per well on a 96-well plate and cultured overnight. The cells were then transfected with 100 ng of a reporter- plasmid (ARE241B sequence reported by Azeem et al. 2017;

subcloned into pTK-Red-firefly (Promega) in front of a CMVmini promoter by M. Hertel) and 2.5 ng of an expression plasmid for the full-length zebra finch androgen receptor (zfAR-FL; sequence custom made by Bio Basic, Canada; subcloned into pcDNA3.12 by Dr. Hertel) or the zebra finch androgen receptor splice variants zfAR-V2-4 and zfAR-V2-8 in pcDNA3.1 (both custom made by Eurofins) using lipofectamine 2000 (Thermo Fisher 11668030) according to manufacturer's instructions. To study the dominant-negative activity of androgen receptor variants HeLa cells were co-transfected with the zfAR-FL and the synthetic human dominant-negative variant ARi410 in pCMV (Butler et al., 2006), its zebra finch homologue (sequence including a C-terminal FLAG-tag synthesised and cloned into pcDNA3.1 by Bio Basic, Canada), zfAR-V2-4 or zfAR-V2-8 at a concentration ratio of 4:1 and subsequently incubated overnight. On the second day the medium was replaced by DMEM high glucose phenol red free (Gibco 31053028) containing 10 % dextran-coated charcoal-stripped FCS (Sigma F6765) and supplemented with 10 nM of 5α-dihydrotestosterone (DHT) or a corresponding volume of the solvent ethanol. Then, cells were incubated overnight, lysed the following day, assayed for luciferase activity using a Luciferase assay kit (P.J.K 102510) and a plate-reading luminometer (Tristar multimode reader LB942, Berthold). To correct for the transfection efficiency in different experiments the data were normalized using the background activity of the reporter plasmid.

Acknowledgments

We are grateful to Anja Lohrentz, Christina Reusch, Judith Kammerlander, for their excellent technical support. IG also thanks the IMPRS research school for training and support. The study was founded by the Max Planck Society.

General discussion

Zebra finch primordial germ cells in culture: an efficient tool to produce transgenic songbirds

Primordial germ cells (PGCs) constitute the key component of the most efficient method to produce transgenic chickens. They can be cultured, genetically modified and clonally selected *in vitro*, before reinjection into host embryos where they will contribute to the germ line and generate transgenic offspring (Van De Lavoir et al. 2006; Macdonald et al. 2010; Schusser et al., 2013, 2016; Dimitrov et al. 2016). In other avian species such as the songbirds the generation of transgenic birds rely on less efficient methods like the genetic manipulation of PGCs *in ovo* to generate transgenic models which has hinder the development of functional genomic studies in these species. In the scope of my thesis I have adapted the method available to generate transgenic chickens by the use of cultured PGCs to produce transgenic songbirds.

Zebra finch PGCs (zfPGCs) were extracted from the blood of HH stage 14 embryos and cultured under our defined culture conditions optimized for zfPGCs. During the first four days *in vitro* (DIV) red blood cells die and zfPGCs start to proliferate. At 7 DIV their genome was genetically modified by the use of lentiviral vectors carrying the enhanced Green Fluorescence Protein (eGFP) gene under control of the human PGK promoter. At this time point zfPGCs were growing in clumps and showed very high transformation efficiency with almost all PGCs expressing the eGFP protein two days after transformation. At 10 DIV the genetically modified zfPGCs were injected into EGK stage X zebra finch embryos. Ten founder birds were raised to sexual maturity and mated with wild type zebra finches. All founders produced transgenic offspring in the next generation with a germline transmission rate between 4 and 22 %. The transgenic zebra finches expressed the transgene in kidney, heart, lung, liver, ovary and brain including the song control system with the percentage of eGFP expressing cells varying from 1 to 70% depending on the organ and the individual.

Several methods were developed for the generation of transgenic birds with variable ranges of efficiency measured in germline transmission rate. Some methods such as the *in ovo* injection of transposons and the sperm transfection assisted gene editing had a low germline transmission rate of about 2 and 7 %, respectively (Tyack et al., 2013; Cooper et al., 2017). The only technique developed until now to produce transgenic zebra finches was the direct injection of viral vectors in the blastoderm, a method that showed a

germline transmission rate of around 18 % in chickens, 7 % in quails and 10 % in zebra finches with only 3 out of 23 founders producing transgenic offspring in the last species (Mizuarai et al. 2001; Mozdziak et al., 2003; McGrew et al. 2004; Zhang et al. 2012; Agate et al., 2009). Since the genetic modification of cultured PGCs is the most efficient method used so far to produce transgenic chickens with a germline transmission rate of on average 20 and up to 95 % (Van De Lavoir et al., 2006; Mac donald et al., 2010; Schusser et al., 2013, 2016; Dimitrov et al., 2016) enables the use of modern gene editing techniques, here I have adapted and applied this method for the generation of transgenic zebra finches. Using this method, I have obtained a germline transmission rate between 4 and 22 %. Additionally, responsible for the high efficiency of this technique compared to the direct viral injection method used previously in zebra finches all founders in this study produced transgenic offspring. In conclusion, the culture and genetic modification of zfPGCs constitute a more efficient method to target zfPGCs that will migrate to the gonads, contribute to the germline and successfully generate transgenic offspring in the next generation.

Proof of principle: generation of transgenic zebra finches expressing eGFP with cultured PGCs

To test the use of cultured zfPGCs that were genetically modified by lentiviral vectors for the generation of a transgenic zebra finches transgenic were developed that expressed eGFP under control of the human PGK-promoter. The expression of the transgene in the transgenic zebra finches was not ubiquitous but restricted to clusters of cells in the brain, including the song control system, kidney, heart, liver, lung and ovary as discovered by anti-GFP immunostainings. The variegated expression pattern of eGFP might have been the result of the down-regulation or silencing of the transgene in most cells and a consequence of epigenetic modifications of the transgene like DNA-methylation and histone deacetylation or chromatin condensation in the vicinity of the integration site (Dobie et al., 1997; Razin 1998; Bird et al., 1999; Ellis 2005; Mok et al.,2007). Such gene expression alterations are commonly observed in transgenic models that use retroviruses as vectors, resulting in the coexistence of cells that express the transgene at a high and low level, respectively, within a given tissue (Yao et al., 2004; Baup et al., 2009).

In follow-up studies a variegated expression might be avoidable by using transgene constructs that include an ubiquitous chromatin opening element (UCOE) as it can be derived form the human HNRPA2B1-CBX3 locus (A2UCOE) and allows position-independent, long-term transgene expression (Williams et al., 2005; Zhang et al,. 2010; Hoffman et al., 2017). The full-length A2UCOE and the minimal CBX3 element can successfully prevent DNA methylation and maintain transgene expression in hematopoietic stem cells, embryonic carcinoma cells and human induced pluripotent stem cells (iPSC) (Zhang et al., 2007, 2010; Hoffman et al., 2017).

Another reason for the uneven expression pattern of eGFP in the transgenic zebra finches might have been the cell type specific activity of the PGK promoter. This promoter is ubiquitously expressed in human cells and caused a high and homogeneous expression level of eGFP in cultured zfPGCs. The PGK promoter has not been tested before in a transgenic zebra finch model and was chosen here because it appeared to be well suited for genome editing approaches. For a high and ubiquitous transgene expression in zebra finches additional promoter constructs need to be tested in the future. In particular, to match the transgene expression pattern to the specific requirements that are defined by the various experimental questions addressed to a transgenic songbird model.

Application of recent genome editing techniques in cultured zfPGCs

The new method presented here for the generation of transgenic songbirds offers the possibility of transgene over-expressions, gene knock-down and knock-out. For example, to study the role of genes related to the development of the song control system such as sex hormone receptors (Frankl-Vilches and Gahr, 2017) and brain derived neurotrophic factor (BDNF) (Dittrich et al. 2013, Tang et al., 2013) as well as genes related to song learning like FoxP2 (Haesler et al 2007; Kosubek-Langerand Scharff 2020). Furthermore, this method could also be extended to other song bird species that can be breed easily allowing for the study of genes relevant to research topics like taste and olfactory receptors in birds or avian mating and aggression behaviour among others.

Comparable to the eGFP zebra finch model generated here transgenic zebra finches could be developed in future with an over-expression of transgenes in the whole body by the

use of constructs that include UCOEs or in particular cells when working with tissue- or cell-type specific promoters. The method could be applied for a conditional mutagenesis approach when over-expressing the genes for Cas9, Cre or Flp. Using such transgenic birds CRISPR/Cas9 approaches can be reduced to the local application of guide RNAs or floxed/FRT cassettes flanking the transgene for which the time point of application can be selected depending on the scientific question (Gossen and Bujard, 1992; Kistnert et al., 1996; McLellanet al., 2017). In order to investigate the function of genes that would result lethal or would produce serious developmental deficiencies when knocked-out, alternatives are the down-regulation of a gene by applying the RNA interference (RNAi) technology or over-expressing dominant negative mutant variants of a gene. RNAi effects can be achieved by incorporating shRNA or dsRNA with lentiviral vectors to reduce the expression of a particular gene by mRNA cleavage the RNA induced silencing complex (RISC) (Chang et al., 2005; Peng et al., 2006; Seibler et al., 2007; Dickinset al., 2007). The disadvantage of siRNA is the low specificity with off-target effects (Jackson and Linsley 2010). An over-expression of a dominant negative mutant gene would be particularly interesting for the study of the effect of sex hormones such as androgens and estrogen in the development of the song control system and behaviour of songbirds (see chapter 3). Finally, knock-out zebra finches can be generated by the application of gene targeting techniques such as CRISPR/Cas9 to zfPGCs in culture using highly efficient lentiviral vectors to induce indel-type mutations by non-homologous end joining that will inactivate the target gene. For more complex genome editing techniques that involve a homology directed repair (HDR), culture conditions that permit a more extended growth of zfPGCs *in vitro* and clonal selection of genetically modified zfPGCs will need to be developed. Additionally, potential of off-target effects need to be carefully investigated (Hsu et al., 2013; Veres et al., 2014; Kim et al., 2016).

Long-term expansion of zfPGCs *in vitro*

In order to apply precise genome editing tools where a donor DNA is provided and homologous directed repair is produced when using CRISPR/Cas9, it seems necessary to culture the PGCs for an extended period of time. Because that allows for the application of gene editing techniques followed by antibiotic selection and clonal expansion of the targeted cells. However, so far only chicken PGCs can be cultured for a period of time

longer than one month (Whyte et al., 2015) so that they could be used for the generation of several chicken transgenic models (Oishi et al. 2016; Zuo et al. 2016; Zhang et al. 2017). In contrast to chicken PGCs, zfPGCs can so far only be cultured for up to 30 days (Jung et al., 2019; Gessara et al., 2021) and in order to obtain a large number of PGCs blood samples from multiple zebra finch embryos had to be pooled leading to sex-mixed cultures. Thus, selection of zfPGC clones and matching of the sex between the genetically modified zfPGCs and the host embryos could not be performed.

From 10 DIV on the growth curve of cultured zfPGCs enters a stationary phase where they stop dividing and begin to die. To determine the causes of cell death in cultured zfPGCs and to improve the culture conditions, we determined the transcriptome of zfPGCs freshly extracted from blood and zfPGCs cultured for 10 DIV. In the course of the bioinformatic analyses first the expression levels of the Telomeres Reverse Transcriptase (TERT) were compared. Telomeres are linear guanine-rich DNA structures at the ends of chromosomes crucial for cellular self-renewal. With subsequent cell divisions telomeres shorten until they eventually induce senescence to avoid chromosomal damage (reviewed by Chan and Blackburn 2004). TERT is the catalytic subunit of the enzyme Telomerase that in conjunction with its telomerase RNA component (TERC) is sufficient to extend the telomeric length. High TERT expression levels are normally found in stem cells and germline cells, enhancing proliferation and maintaining pluripotency (Yang et al., 2008; Xie et al., 2011). In zfPGCs the TERT expression was found unaltered in the transcriptome of zfPGCs cultured for 10 DIV when compared to freshly extracted zfPGCs.

In contrast, transcriptome comparisons revealed pathways related to oxidoreductase activity acting on peroxidase as acceptor to be down-regulated in cells cultured for 10 days. Since this result could indicate an inadequate supply of antioxidants in the culture medium, and since oxidative stress constitutes a cause of cell death addition of antioxidant supplements to the culture medium was tested but found to have no significant effect on the survival time of zfPGCs *in vitro*. Also, expression of key apoptotic genes was found in the transcriptome to be unaltered or down-regulated in zfPGCs cultured for 10 DIV. Disregarding the hypothesis that apoptosis is the main cause for the cultured zfPGCs to enter the stationary phase of their growth curve.

Further, pathways were identified that are regulated by (Matsumoto et al., 2001; Scaltriti et al., 2006) and VEGF (Nalbandian et al., 2003; Celik-Ozenci et al., 2003; Tian et al., 2016) and were altered in the transcriptome of zfPGCs at 10 DIV. Therefore, the effects of an addition of EGF and VEGF to the culture medium on the growth of zfPGCs at 15 DIV was tested. There was a significant increase in the cell number of zfPGCs cultured in the presence of VEGF and a positive trend in cultures containing EGF when compared to control cultures that did not contain any of these growth factors.

Other growth factors that stimulate pathways related to cell proliferation and cell survival and that could contribute to the proliferation of zfPGCs *in vitro*, were reported to be LIF (Resnick et al., 1992; Makoolati et al., 2016), TGF-ß and SCF (Weiss et al., 2012, Graf et al., 2011; Sette et al. 2000; De Felici 2000; Liu et al. 2007). In particular, LIF apparat from stimulating cell proliferation *in vivo*, also contributes to the maintenance of stem cells in an undifferentiated state (Lavial et al., 2007). While there was no effect on the proliferation of cultured zfPGCs in the presence of TGF-ß or SCF a significant increase in the zfPGCs number in cultures containing LIF was observable.

In conclusion, incorporation of the the growth factors VEGF, EGF and LIF led to an increase in the number of zfPGCs available for a gene editing process in a short period of time in culture and injection into a host embryo. However, the survival time of zfPGCs *in vitro* could not be extended by these growth factors.

Future directions to improve the culture conditions of zfPGCs

Chickens constitute the only avian species in which PGCs can be expanded for an indefinite period of time conserving their undifferentiated state. Therefore, a comparison of the zebra finch and the chicken PGC transcriptome might help to identify key genes or pathways that contribute to the maintenance of zfPGC renewal *in vitro*.

Although mRNA sequencing constitutes an extremely useful tool to identify differentially expressed genes and determine the gene networks associated with specific physiological or pathological processes, a substantial part of the information about gene expression might be lost in different steps of the bioinformatic analysis process when studying species other than the commonly used model organisms. Particularly, when the reads were aligned to the reference genome of the zebra finch in the mapping stage, about 50 %

of the sequence information was lost with the reads that could not being mapped. A second major loss of information occurred when the orthologous human genes had to be identified for the functional analysis. For several zebra finch genes human orthologs could not being assigned so that they not considered in downstream analyses. Therefore, the zebra finch reference genome needs to be further improved in order to discover more changes in the physiology of zfPGCs evoked by the culture conditions.

After 4 DIV, zfPGCs start forming loose cellular aggregations which was beneficial for a highly efficient transduction with lentiviral vectors when compared to singly growing cells. It is still unclear whether the formation of aggregations facilitated the genetic modification at the expense of the long term survival of the zfPGCs *in vitro*. However, when the cellular aggregations were dissociated with papain cytoplasmic vesicles reappeared and the survival time of the zfPGCs *in vitro* did not increase. Since zfPGCs migrate though the blood system *in ovo*, the effects of a feeder layer formed by endothelial cells on the survival and transduction efficiency of cultured zfPGCs should be tested in future.

To expand the method that was developed in this thesis for the generation of transgenic zebra finches to other songbird species several requirements need to be fulfilled: First of all, continuous breeding of the birds is needed for the production of larger numbers of eggs to isolate PGCs and to obtain host embryos at the required developmental stage for the injection of the genetically modified PGCs. Next, synchrony of foster parent egg laying and PGC growth in vitro is required. The number of PGCs that could be isolated from a zebra finch egg varied strongly between individual breeding couples, so that they could be selected to deliver eggs with a high or low PGC content to be used as donors and recipients, respectively. Further, lentiviral vectors turned out to be highly efficient for the transduction and genetic modification of cultured zfPGCs under the optimized culture conditions. Whether all these findings hold true for other songbirds species still need to be specified.

Next steps in the generation of a transgenic zebra finch models

The genetic modification of the androgen receptor (AR) constitutes an excellent approach to study the role of androgens such as testosterone and DHT in the development of the

song control system in birds. Therefore, the availability of an AR knock-down zebra finch would be of great advantage. One of the zebra finch AR splice variants studied in this project, the zfAR-V2-8 showed an inhibitory effect on the activity of the zfAR-FL when co-transfected *in vitro* at a ratio of 4:1 in an AR luciferase reporter assay. This naturally occurring AR variant could reduce the activity of the zfAR-FL by up to 55 % so that the zfAR-V2-8 constitutes an appropriate candidate transgene to be used in the generation of a transgenic AR knock- down zebra finch by the means of a dominant-negative approach. Since this variant was found to be widely expressed naturally in different tissues, the over-expression of zfAR-V2-8 is not expected to produce pathological side effects. To obtain an even larger reduction of the zfAR-FL activity the synthetic human dominant negative variant ARi410 (Butler et al., 2006) or its zebra finch homologue zfARi410 could be used as an over-expressed transgene. With these two synthetic AR variants I detected a reduction in the transactivation effect that was 12 to 17 % larger than that obtained with zfAR-V2-8. However, silencing of transgene expression delivered by lentiviral vectors will make it necessary to improve the vector design for a dominant-negative effect of the transgene *in vivo*.

Previous studies developed in mouse models (Yeh et al., 2002; Sato et al., 2004) determined that a knock-out of the androgen receptor does not result lethal, and that androgens are necessary for perinatal masculinization of the male brain and to induce male-typical behaviours in the adult (Sato et al., 2004). The availability of a transgenic AR knock-out zebra finch model would open the possibility to study the effect of androgens from the early embryonic development on. For studies of the effects of maternal androgens delivered to the egg (Adkins-Regan et al., 2013; Tschirren 2015) a knock-down model would be a new approach (Podmokla et al., 2018) to investigate the role of prenatal androgens for the early bird development. To schedule the onset of knock-down effects to particular post-natal developmental stages, the song learning phase or particular nuclei of the song control system PGCs could be transduced with dominant negative AR variant transgenes designed with flanking loxP sites followed by the application of a CRE enzyme at the desired time point and location.

In this study the AR variants zfAR-V2-4 and zfAR-V2-8 were found to be expressed in HVC at 30, 50 and 100 days post hatching. To understand their role in HVC during development and the song learning process, additional studies could be done. First, cell type specificity of the expression of each variant in HVC can be identified, for example

by using novel in situ hybridization techniques like RNAscope where zfAR variant specific probes might be designed for the transcript specific UTR-sequences combined with labelling of projecting neurons and probes that detect cell-type specific markers like GAD65/67 in interneurons. To gain information about the function of the zfAR variants in HVC during the song development process viral vectors for Cas9 and zfAR variant specific guide RNAs can be designed for a local knock-down . A combination of local injection of one of the two CrispR/Cas9 components into HVC and retrograde delivery of the other at a selected developmental time point can be performed to target HVC projection neurons.

Since the effects of the zfAR variants (zfAR-FL, zfAR-V2-4 and zfAR-V2-8) dependent on their relative protein concentrations techniques should be established to measure them in different tissues and at different developmental stages. With antibody-based methods that allow a discrimination of the zfAR variants like Western-blot analysis one could determine how they balance each other. Next to the zebra finch this should be done in avian species where the development of the song control system and singing behaviour is strongly testosterone-dependent such as the canary (Alward et al., 2018). A study in this bird species that determines the seasonal variations in the protein levels of zfAR-FL, zfAR-V2-4 and zfAR-V2-8 and correlates these data with variations in the seasonal androgen levels would bring new insights into the function of the AR variants for seasonal changes in the singing behaviour.

When testing the zfAR-V2-4 in an in vitro AR promoter assay, both constitutive and ligand dependent transactivational activity were observed. Enhanced activity in the presence of the ligand suggested the presence of a ligand binding site in zfAR-V2-4 different from that found in zfAR-FL where it is encoded by exons 5-8. To provide evidences that the zfAR-V2-4 is effectively binding to androgens, additional tests are necessary. A potential androgen binding site might be encoded by the new 59 c-terminal end amino acids of the variant. Mutants should be designed that carry amino acid exchanges or deletions within this sequence and tested in the AR promoter assay. Further, to demonstrate androgen binding directly, the zfAR-V2-4 could be incubated with radioactive labelled testosterone and analysed in a dot plot assay. The zfAR-V2-8 was ubiquitously expressed in all the tissues tested and presented an inhibitory effect over the zfAR-FL when tested together in an in vitro AR promoter assay. This indicates that the function of the zfAR-V2-8 could consist in limiting the effect of androgens in the body.

Interestingly, another natural occurring AR variant was discovered in humans (hAR-45) and was first described to have an inhibitory effect on the hAR-FL activity (Ahren-Fath et al., 2005). Later, hAR-45 was identified as the membrane-associated AR in the brain (Garza-Conteras et al., 2017; Schreihofer et al., 2018) responsible for fast, non-genomic signal transduction of androgen activity in the brain (Wu et al., 2008; Garza-Conteras et al., 2017; Duong et al., 2020). Since the exon composition of hAR-45 and zfAR-V2-8 are comparable it is tempting to speculate that the zfAR-V2-8 exerts cell membrane signal transduction of androgens as well. In contrast to the zfAR-V2-8 that showed little developmental expression changes in the postnatal brain the zfAR-FL was stronger expressed in more mature brain tissues where the song control system is sexually dimorph. If it is the genomic activity of zfAR variants that affects the sexual differentiation of the brain the use of the zfAR-V2-8 in a dominant-negative transgenic approach can be expected to lead to new insights into the role of androgens in the sexual differentiation of the song control system.

References

Abe K, Matsui S, Watanabe D (2015) Transgenic songbirds with suppressed or enhanced activity of CREB transcription factor. *Proc Natl Acad Sci U S A* 112(24):7599–7604.

Adam I, Mendoza E, Kobalz U, Wohlgemuth S, Scharff C (2016) FoxP2 directly regulates the reelin receptor VLDLR developmentally and by singing. *Molecular and Cellular Neuroscience*, 74: 96-105.

Afgan E, Baker D, Batut B, Van den Beek M, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Grüning B (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update, *Nucleic Acids Research* 46 (W1): W537–W544.

Adkins-Regan, E., Banerjee, S. B., Correa, S. M., & Schweitzer, C. (2013). Maternal effects in quail and zebra finches: behavior and hormones. General and Comparative Endocrinology, 190, 34-41.

Adret, P., and Margoliash, D. (2002). Metabolic and neural activity in the song system nucleus robustus archistriatalis: effect of age and gender. The Journal of comparative neurology, 454(4), 409–423.

Agate, R. J., Choe, M., and Arnold, A. P. (2004). Sex differences in structure and expression of the sex chromosome genes CHD1Z and CHD1W in zebra finches. Molecular biology and evolution, 21(2), 384–396.

Agate, R. J., Grisham, W., Wade, J., Mann, S., Wingfield, J., Schanen, C., Palotie, A., and Arnold, A. P. (2003). Neural, not gonadal, origin of brain sex differences in a gynandromorphic finch. Proceedings of the National Academy of Sciences of the United States of America, 100(8), 4873–4878.

Agate RJ, Scott BB, Haripal B, Lois C, Nottebohm F (2009) Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning. *Proc Natl Acad Sci U S A* 106(42):17963–17967.

Ahmadiantehrani, S., and London, S. E. (2017). A reliable and flexible gene manipulation strategy in posthatch zebra finch brain. Scientific reports, 7, 43244.

Ahrens-Fath, I., Politz, O., Geserick, C., and Haendler, B. (2005). Androgen receptor

function is modulated by the tissue-specific AR45 variant. The FEBS journal, 272(1), 74–84.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008). Molecular biology of the cell. New York: Garland Science.

Alward, B. A., Balthazart, J., and Ball, G. F. (2013). Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate. Proceedings of the National Academy of Sciences of the United States of America, 110(48), 19573–19578.

Alward, B. A., Madison, F. N., Parker, S. E., Balthazart, J., and Ball, G. F. (2016). Pleiotropic Control by Testosterone of a Learned Vocal Behavior and Its Underlying Neuroplasticity(1,2,3). eNeuro, 3(1), ENEURO.0145-15.2016.

Alward, B. A., Cornil, C. A., Balthazart, J., & Ball, G. F. (2018). The regulation of birdsong by testosterone: Multiple time-scales and multiple sites of action. Hormones and behavior, 104, 32-40.

Amendt, C., Schirmacher, P., Weber, H., and Blessing, M. (1998). Expression of a dominant negative type II TGF-beta receptor in mouse skin results in an increase in carcinoma incidence and an acceleration of carcinoma development. Oncogene, 17(1), 25–34.

Apfelbeck, B., Kiefer, S., Mortega, K. G., Goymann, W., and Kipper, S. (2012). Testosterone affects song modulation during simulated territorial intrusions in male black redstarts (Phoenicurus ochruros). PloS one, 7(12), e52009.

Arnold A. P. (1975). The effects of castration on song development in zebra finches (Poephila guttata). The Journal of experimental zoology, 191(2), 261–278.

Aronov, D., Andalman, A. S., and Fee, M. S. (2008). A specialized forebrain circuit for vocal babbling in the juvenile songbird. Science (New York, N.Y.), 320(5876), 630–634.

Balthazart, J., Foidart, A., Wilson, E. M., and Ball, G. F. (1992). Immunocytochemical localization of androgen receptors in the male songbird and quail brain. The Journal of comparative neurology, 317(4), 407–420.

Balthazart J, Charlier TD, Barker JM, Yamamura T, Ball GF (2010) Sex steroid-induced neuroplasticity and behavioral activation in birds. Eur J Neurosci 32(12):2116–2132.

Bender, A. T., and Veney, S. L. (2008). Treatment with the specific estrogen receptor antagonist ICI 182,780 demasculinizes neuron soma size in the developing zebra finch brain. Brain research, 1246, 47–53.

Bernard, D. J., Bentley, G. E., Balthazart, J., Turek, F. W., and Ball, G. F. (1999). Androgen receptor, estrogen receptor alpha, and estrogen receptor beta show distinct patterns of expression in forebrain song control nuclei of European starlings. Endocrinology, 140(10), 4633–4643.

Bevan, C. L., Hoare, S., Claessens, F., Heery, D. M., and Parker, M. G. (1999). The AF1 and AF2 domains of the androgen receptor interact with distinct regions of SRC1. Molecular and cellular biology, 19(12), 8383–8392.

Birkhead, T. R., Pellatt, J., and Hunter, F. M. (1988). Extra-pair copulation and sperm competition in the zebra finch. Nature, 334(6177), 60–62.

Bockamp, E., Sprengel, R., Eshkind, L., Lehmann, T., Braun, J. M., Emmrich, F., and Hengstler, J. G. (2008). Conditional transgenic mouse models: from the basics to genome-wide sets of knockouts and current studies of tissue regeneration. Regenerative medicine, 3(2), 217–235.

Böhner, J., Chaiken, M. L., Ball, G. F., and Marler, P. (1990). Song acquisition in photosensitive and photorefractory male European starlings. Hormones and behavior, 24(4), 582–594.

Bolger A, Lohse M, Usadel B, (2014) Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* 30 (15): 2114-2120, https://doi.org/10.1093/bioinformatics/btu170

Bolhuis J.J., Gahr M (2006) Song system Neural mechanisms of birdsong memory. *Nat Rev Neurosci* 7(5):347–357

Bolhuis, J. J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. Nature Reviews Neuroscience, 11(11), 747-759.

Bottjer, S. W., and Hewer, S. J. (1992). Castration and antisteroid treatment impair vocal learning in male zebra finches. Journal of neurobiology, 23(4), 337–353.

Böttcher, R. T., and Niehrs, C. (2005). Fibroblast growth factor signaling during early

vertebrate development. Endocrine reviews, 26(1), 63–77.

Bottjer, S. W., Glaessner, S. L., and Arnold, A. P. (1985). Ontogeny of brain nuclei controlling song learning and behavior in zebra finches. The Journal of neuroscience : the official journal of the Society for Neuroscience, 5(6), 1556–1562.

Bouabe, H., and Okkenhaug, K. (2013). Gene targeting in mice: a review. Methods in molecular biology (Clifton, N.J.), 1064, 315–336.

Brandl, H. B., Griffith, S. C., Farine, D. R., and Schuett, W. (2021). Wild zebra finches that nest synchronously have long-term stable social ties. The Journal of animal ecology, 90(1), 76–86.

Breier, G., Albrecht, U., Sterrer, S., and Risau, W. (1992). Expression of vascular endothelial growth factor during embryonic angiogenesis and endothelial cell differentiation. Development (Cambridge, England), 114(2), 521–532.

Breier, G., Clauss, M., and Risau, W. (1995). Coordinate expression of vascular endothelial growth factor receptor-1 (flt-1) and its ligand suggests a paracrine regulation of murine vascular development. Developmental dynamics : an official publication of the American Association of Anatomists, 204(3), 228–239.

Buehr M. (1997). The primordial germ cells of mammals: some current perspectives. Experimental cell research, 232(2), 194–207.

Butler, L. M., Centenera, M. M., Neufing, P. J., Buchanan, G., Choong, C. S., Ricciardelli, C., Saint, K., Lee, M., Ochnik, A., Yang, M., Brown, M. P., and Tilley, W. D. (2006). Suppression of androgen receptor signaling in prostate cancer cells by an inhibitory receptor variant. Molecular endocrinology (Baltimore, Md.), 20(5), 1009–1024.

Callewaert, L., Van Tilborgh, N., and Claessens, F. (2006). Interplay between two hormone-independent activation domains in the androgen receptor. Cancer research, 66(1), 543–553.

Cao, S., Zhan, Y., and Dong, Y. (2016). Emerging data on androgen receptor splice variants in prostate cancer. Endocrine-related cancer, 23(12), T199–T210.

Capecchi M. R. (1989). Altering the genome by homologous recombination. Science (New York, N.Y.), 244(4910), 1288–1292.

Carbery, I. D., Ji, D., Harrington, A., Brown, V., Weinstein, E. J., Liaw, L., and Cui, X. (2010). Targeted genome modification in mice using zinc-finger nucleases. Genetics, 186(2), 451–459.

Carlson, D. F., Tan, W., Lillico, S. G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D. F., Long, C. R., Whitelaw, C. B., and Fahrenkrug, S. C. (2012). Efficient TALEN-mediated gene knockout in livestock. Proceedings of the National Academy of Sciences of the United States of America, 109(43), 17382–17387.

Carrer, H. F., and Cambiasso, M. J. (2002). Sexual differentiation of the brain: genes, estrogen, and neurotrophic factors. Cellular and molecular neurobiology, 22(5-6), 479–500.

Carsience, R. S., Clark, M. E., Verrinder Gibbins, A. M., and Etches, R. J. (1993). Germline chimeric chickens from dispersed donor blastodermal cells and compromised recipient embryos. Development (Cambridge, England), 117(2), 669–675.

Celik-Ozenci, C., Akkoyunlu, G., Kayisli, U. A., Arici, A., and Demir, R. (2003). Localization of vascular endothelial growth factor in the zona pellucida of developing ovarian follicles in the rat: a possible role in destiny of follicles. Histochemistry and cell biology, 120(5), 383–390.

Céraline, J., Erdmann, E., Erbs, P., Deslandres-Cruchant, M., Jacqmin, D., Duclos, B., Klein-Soyer, C., Dufour, P., and Bergerat, J. P. (2003). A yeast-based functional assay for the detection of the mutant androgen receptor in prostate cancer. European journal of endocrinology, 148(1), 99–110.

Chang, L. J., Liu, X., and He, J. (2005). Lentiviral siRNAs targeting multiple highly conserved RNA sequences of human immunodeficiency virus type 1. Gene therapy, 12(14), 1133–1144.

Chen, Z., Ye, R., and Goldman, S. A. (2013). Testosterone modulation of angiogenesis and neurogenesis in the adult songbird brain. Neuroscience, 239, 139–148.

Cho, S. W., Kim, S., Kim, J. M., and Kim, J. S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nature biotechnology, 31(3), 230–232.

Cynx, J., and Nottebohm, F. (1992). Role of gender, season, and familiarity in

discrimination of conspecific song by zebra finches (Taeniopygia guttata). Proceedings of the National Academy of Sciences of the United States of America, 89(4), 1368–1371.

Cohen G. M. (1997). Caspases: the executioners of apoptosis. The Biochemical journal, 326 (Pt 1)(Pt 1), 1–16.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science (New York, N.Y.), 339(6121), 819–823.

Cooke B. M. (2006). Steroid-dependent plasticity in the medial amygdala. Neuroscience, 138(3), 997–1005.

Cooper, C. A., Challagulla, A., Jenkins, K. A., Wise, T. G., O'Neil, T. E., Morris, K. R., Tizard, M. L., and Doran, T. J. (2017). Generation of gene edited birds in one generation using sperm transfection assisted gene editing (STAGE). Transgenic research, 26(3), 331–347.

Cordes, M. A., Stevenson, S. A., Driessen, T. M., Eisinger, B. E., and Riters, L. V. (2015). Sexually-motivated song is predicted by androgen-and opioid-related gene expression in the medial preoptic nucleus of male European starlings (Sturnus vulgaris). Behavioural brain research, 278, 12–20.

Cynx, J., Bean, N. J., and Rossman, I. (2005). Testosterone implants alter the frequency range of zebra finch songs. Hormones and behavior, 47(4), 446–451.

D'Costa S, Petitte JN (1999) Characterization of stage-specific embryonic antigen-1 (SSEA-1) expression during early development of the turkey embryo. *Int J Dev Biol* 43(4):349–356.

De Felici M. (2000). Regulation of primordial germ cell development in the mouse. The International journal of developmental biology, 44(6), 575–580.

Dehm, S. M., Schmidt, L. J., Heemers, H. V., Vessella, R. L., and Tindall, D. J. (2008). Splicing of a novel androgen receptor exon generates a constitutively active androgen receptor that mediates prostate cancer therapy resistance. Cancer research, 68(13), 5469–5477.

Denayer, S., Helsen, C., Thorrez, L., Haelens, A., and Claessens, F. (2010). The rules of DNA recognition by the androgen receptor. Molecular endocrinology (Baltimore, Md.),

24(5), 898–913.

Derégnaucourt, S., and Gahr, M. (2013). Horizontal transmission of the father's song in the zebra finch (Taeniopygia guttata). Biology letters, 9(4), 20130247.

Derégnaucourt, S., Mitra, P. P., Fehér, O., Pytte, C., and Tchernichovski, O. (2005). How sleep affects the developmental learning of bird song. Nature, 433(7027), 710–716.

DeVoogd, T. J., and Nottebohm, F. (1981). Sex differences in dendritic morphology of a song control nucleus in the canary: a quantitative Golgi study. The Journal of comparative neurology, 196(2), 309–316.

Dhup, S., and Majumdar, S. S. (2008). Transgenesis via permanent integration of genes in repopulating spermatogonial cells *in vivo*. Nature methods, 5(7), 601–603.

Dickins, R. A., McJunkin, K., Hernando, E., Premsrirut, P. K., Krizhanovsky, V., Burgess, D. J., Kim, S. Y., Cordon-Cardo, C., Zender, L., Hannon, G. J., and Lowe, S. W. (2007). Tissue-specific and reversible RNA interference in transgenic mice. Nature genetics, 39(7), 914–921.

Dimitrov L, Pedersen D, Ching KH, Yi H, Collarini EJ, Izquierdo S, Van de Lavoir MC, Leighton PA (2016) Germline Gene Editing in Chickens by Efficient CRISPR-Mediated Homologous Recombination in Primordial Germ Cells. *PLoS One* 11(4): e0154303 DOI:

Ding, D., Xu, L., Menon, M., Reddy, G. P., and Barrack, E. R. (2004). Effect of a short CAG (glutamine) repeat on human androgen receptor function. The Prostate, 58(1), 23–32.

Ding, D., Xu, L., Menon, M., Reddy, G. P., and Barrack, E. R. (2005). Effect of GGC (glycine) repeat length polymorphism in the human androgen receptor on androgen action. The Prostate, 62(2), 133–139.

Dittrich, F., Feng, Y., Metzdorf, R., and Gahr, M. (1999). Estrogen-inducible, sex-specific expression of brain-derived neurotrophic factor mRNA in a forebrain song control nucleus of the juvenile zebra finch. Proceedings of the National Academy of Sciences of the United States of America, 96(14), 8241–8246.

Dobin A, Davis C, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras T (2013) STAR: ultrafast universal RNA-seq aligner, *Bioinformatics* 29 (1): 15–21

Dolci, S., Williams, D. E., Ernst, M. K., Resnick, J. L., Brannan, C. I., Lock, L. F., Lyman, S. D., Boswell, H. S., and Donovan, P. J. (1991). Requirement for mast cell growth factor for primordial germ cell survival in culture. Nature, 352(6338), 809–811.

Doupe, A. J., Solis, M. M., Kimpo, R., and Boettiger, C. A. (2004). Cellular, circuit, and synaptic mechanisms in song learning. Annals of the New York Academy of Sciences, 1016, 495–523.

Duan J, Duan J, Zhang Z, Tong T. Irreversible cellular senescence induced by prolonged exposure to H2O2 involves DNA-damage-and-repair genes and telomere shortening. Int J Biochem Cell Biol. 2005 Jul;37(7):1407-20.

Dumont, D. J., Jussila, L., Taipale, J., Lymboussaki, A., Mustonen, T., Pajusola, K., Breitman, M., and Alitalo, K. (1998). Cardiovascular failure in mouse embryos deficient in VEGF receptor-3. Science (New York, N.Y.), 282(5390), 946–949.

Duong, P., Tenkorang, M., Trieu, J., McCuiston, C., Rybalchenko, N., and Cunningham, R. L. (2020). Neuroprotective and neurotoxic outcomes of androgens and estrogens in an oxidative stress environment. Biology of sex differences, 11(1), 12.

Düring, D. N., Dittrich, F., Diales da Rocha, M., Tachibana, R. O., Mori, C., Okanoya, K., et al. (2020). Fast retrograde access to projection neuron circuits underlying vocal learning in songbirds. Cell Reports, 33(6): 108364.

Elmore S. (2007). Apoptosis: a review of programmed cell death. Toxicologic pathology, 35(4), 495–516.

Ergün, S., Kiliç, N., Fiedler, W., and Mukhopadhyay, A. K. (1997). Vascular endothelial growth factor and its receptors in normal human testicular tissue. Molecular and cellular endocrinology, 131(1), 9–20.

Estrada, M., Espinosa, A., Müller, M., and Jaimovich, E. (2003). Testosterone stimulates intracellular calcium release and mitogen-activated protein kinases via a G protein-coupled receptor in skeletal muscle cells. Endocrinology, 144(8), 3586–3597.

Evans, M. J., Carlton, M. B., and Russ, A. P. (1997). Gene trapping and functional genomics. Trends in genetics : TIG, 13(9), 370–374.

Ewerling, S., Hofmann, A., Klose, R., Weppert, M., Brem, G., Rink, K., Pfeifer, A., and Wolf, E. (2006). Evaluation of laser-assisted lentiviral transgenesis in bovine. Transgenic

research, 15(4), 447–454.

Eyal-Giladi H, Kochav S, Menashi MK (1976) On the Origin of Primordial Germ Cells in the Chick Embryo. *Differentiation* 6(1–3):13–16.

Fee, M. S., and Scharff, C. (2010). The songbird as a model for the generation and learning of complex sequential behaviors. ILAR journal, 51(4), 362–377.

Ferro, P., Catalano, M. G., Dell'Eva, R., Fortunati, N., and Pfeffer, U. (2002). The androgen receptor CAG repeat: a modifier of carcinogenesis?. Molecular and cellular endocrinology, 193(1-2), 109–120.

Finkelshtein D, Werman A, Novick D, Barak S, Rubinstein M. (2013) LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. Proc Natl Acad Sci U S A.;110(18):7306-7311.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature, 391(6669), 806–811.

Flamme, I., Breier, G., and Risau, W. (1995). Vascular endothelial growth factor (VEGF) and VEGF receptor 2 (flk-1) are expressed during vasculogenesis and vascular differentiation in the quail embryo. Developmental biology, 169(2), 699–712.

Fortune, E. S., and Margoliash, D. (1995). Parallel pathways and convergence onto HVc and adjacent neostriatum of adult zebra finches (Taeniopygia guttata). The Journal of comparative neurology, 360(3), 413–441.

Frankl-Vilches, C., Kuhl, H., Werber, M., Klages, S., Kerick, M., Bakker, A., de Oliveira, E. H., Reusch, C., Capuano, F., Vowinckel, J., Leitner, S., Ralser, M., Timmermann, B., and Gahr, M. (2015). Using the canary genome to decipher the evolution of hormone-sensitive gene regulation in seasonal singing birds. Genome biology, 16(1), 19.

Frankl-Vilches, C., and Gahr, M. (2018). Androgen and estrogen sensitivity of bird song: a comparative view on gene regulatory levels. Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology, 204(1), 113–126.

Fusani, L., Van't Hof, T., Hutchison, J. B., and Gahr, M. (2000). Seasonal expression of androgen receptors, estrogen receptors, and aromatase in the canary brain in relation to circulating androgens and estrogens. Journal of neurobiology, 43(3), 254–268.

Fujimoto T, Ukeshima A, Kiyofuji R (1976) The origin, migration and morphology of the primordial germ cells in the chick embryo. *Anat Rec* 185(2):139–153.

Gahr, M., and Konishi, M. (1988). Developmental changes in estrogen-sensitive neurons in the forebrain of the zebra finch. Proceedings of the National Academy of Sciences of the United States of America, 85(19), 7380–7383.

Gahr, M., and Metzdorf, R. (1997). Distribution and dynamics in the expression of androgen and estrogen receptors in vocal control systems of songbirds. Brain research bulletin, 44(4), 509–517.

Gahr, M., and Metzdorf, R. (1999). The sexually dimorphic expression of androgen receptors in the song nucleus hyperstriatalis ventrale pars caudale of the zebra finch develops independently of gonadal steroids. The Journal of Neuroscience, 19(7), 2628-2636.

Gahr, M., Metzdorf, R., and Aschenbrenner, S. (1996). The ontogeny of the canary HVC revealed by the expression of androgen and oestrogen receptors. Neuroreport, 8(1), 311–315.

Gahr M. (2004). Hormone-dependent neural plasticity in the juvenile and adult song system: what makes a successful male?. Annals of the New York Academy of Sciences, 1016, 684–703.

Gahr, M., Metzdorf, R., Schmidl, D., and Wickler, W. (2008). Bi-directional sexual dimorphisms of the song control nucleus HVC in a songbird with unison song. PloS one, 3(8), e3073.

Gahr, M., and Konishi, M. (1988). Developmental changes in estrogen-sensitive neurons in the forebrain of the zebra finch. Proceedings of the National Academy of Sciences of the United States of America, 85(19), 7380–7383.

Gahr, M., Güttinger, H. R., and Kroodsma, D. E. (1993). Estrogen receptors in the avian brain: survey reveals general distribution and forebrain areas unique to songbirds. The Journal of comparative neurology, 327(1), 112–122.

Gahr, M., Sonnenschein, E., and Wickler, W. (1998). Sex difference in the size of the neural song control regions in a dueting songbird with similar song repertoire size of males and females. The Journal of neuroscience : the official journal of the Society for

Neuroscience, 18(3), 1124–1131.

Gahr, M., Flügge, G., and Güttinger, H. R. (1987). Immunocytochemical localization of estrogen-binding neurons in the songbird brain. Brain research, 402(1), 173–177.

Gahr M. (1990). Delineation of a brain nucleus: comparisons of cytochemical, hodological, and cytoarchitectural views of the song control nucleus HVc of the adult canary. The Journal of comparative neurology, 294(1), 30–36.

Gahr M. (2003). Male Japanese quails with female brains do not show male sexual behaviors. Proceedings of the National Academy of Sciences of the United States of America, 100(13), 7959–7964.

Gahr M (2007) Sexual Differentiation of the Vocal Control System of Birds. *Adv Genet* 59:67–105. DOI:

Gaj, T., Gersbach, C. A., and Barbas, C. F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology, 31(7), 397–405.

Garelick, T., and Swann, J. (2014). Testosterone regulates the density of dendritic spines in the male preoptic area. Hormones and behavior, 65(3), 249–253.

Garza-Contreras, J., Duong, P., Snyder, B. D., Schreihofer, D. A., and Cunningham, R. L. (2017). Presence of Androgen Receptor Variant in Neuronal Lipid Rafts. eNeuro, 4(4), ENEURO.0109-17.2017.

Gerber, H. P., Malik, A. K., Solar, G. P., Sherman, D., Liang, X. H., Meng, G., Hong, K., Marsters, J. C., and Ferrara, N. (2002). VEGF regulates haematopoietic stem cell survival by an internal autocrine loop mechanism. Nature, 417(6892), 954–958.

Highly efficient genome modification of cultured primordial germ cells with lentiviral vectors to generate transgenic songbirds

Gessara, I., Dittrich, F., Hertel, M., Hildebrand, S., Pfeifer, A., Frankl-Vilches, C., McGrew, M., Gahr, M. Highly efficient genome modification of cultured primordial germ cells with lentiviral vectors to generate transgenic songbirds. bioRxiv 2020.08.11.245753; doi:

Geurts, A. M., Cost, G. J., Freyvert, Y., Zeitler, B., Miller, J. C., Choi, V. M., Jenkins, S. S., Wood, A., Cui, X., Meng, X., Vincent, A., Lam, S., Michalkiewicz, M., Schilling, R.,

Foeckler, J., Kalloway, S., Weiler, H., Ménoret, S., Anegon, I., Davis, G. D., … Buelow, R. (2009). Knockout rats via embryo microinjection of zinc-finger nucleases. Science (New York, N.Y.), 325(5939), 433.

Gill F, Donsker D, editors. (2014). IOC World Bird List (v 4.1). doi: 10.14344/IOC.ML.4.1. Available at: http://www.worldbirdnames.org/

Gill, A., Jamnongjit, M., and Hammes, S. R. (2004). Androgens promote maturation and signaling in mouse oocytes independent of transcription: a release of inhibition model for mammalian oocyte meiosis. Molecular endocrinology (Baltimore, Md.), 18(1), 97–104.

Ginsburg M, Eyal-Giladi H (1986) Temporal and spatial aspects of the gradual migration of primordial germ cells from the epiblast into the germinal crescent in the avian embryo. *Development* 95:53–71. PMID: 3794593

Godin, I., and Wylie, C. C. (1991). TGF beta 1 inhibits proliferation and has a chemotropic effect on mouse primordial germ cells in culture. Development (Cambridge, England), 113(4), 1451–1457.

Goldman SA, Nottebohm F (1983) Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. *Proc Natl Acad Sci U S A* 80(8):2390–2394.

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., and Ruddle, F. H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. Proceedings of the National Academy of Sciences of the United States of America, 77(12), 7380–7384.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proceedings of the National Academy of Sciences of the United States of America, 89(12), 5547–5551.

Graf, U., Casanova, E. A., and Cinelli, P. (2011). The Role of the Leukemia Inhibitory Factor (LIF) - Pathway in Derivation and Maintenance of Murine Pluripotent Stem Cells. Genes, 2(1), 280–297.

Grino, P. B., Griffin, J. E., and Wilson, J. D. (1990). Testosterone at high concentrations interacts with the human androgen receptor similarly to dihydrotestosterone. Endocrinology, 126(2), 1165–1172.

Grisham, W., Park, S. H., Hsia, J. K., Kim, C., Leung, M. C., Kim, L., and Arnold, A. P. (2007). Effects of long-term flutamide treatment during development in zebra finches. Neuroscience letters, 418(1), 92–96.

Gurney, M. E., and Konishi, M. (1980). Hormone-induced sexual differentiation of brain and behavior in zebra finches. Science (New York, N.Y.), 208(4450), 1380–1383.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., and Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. PLoS biology, 5(12), e321.

Hahnloser, R. H., Kozhevnikov, A. A., and Fee, M. S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature, 419(6902), 65–70.

Haile, S., and Sadar, M. D. (2011). Androgen receptor and its splice variants in prostate cancer. Cellular and molecular life sciences : CMLS, 68(24), 3971–3981.

Hamburger V, Hamilton HL (1951) A series of normal stages in the development of the chick embryo. Dev Dyn 88(1):49–92.

Hamson, D. K., Wainwright, S. R., Taylor, J. R., Jones, B. A., Watson, N. V., and Galea, L. A. (2013). Androgens increase survival of adult-born neurons in the dentate gyrus by an androgen receptor-dependent mechanism in male rats. Endocrinology, 154(9), 3294–3304.

Han J.Y., Seo D.S., Shoffner R.N., Guise K.S. (1994) Manipulation of Chicken Primordial Germ Cells for Transgenic Chicken. In: Kobayashi T., Kitagawa Y., Okumura K. (eds) Animal Cell Technology: Basic and Applied Aspects. The Sixth International Meeting of Japanese Association for Animal Cell Technology JAACT'93, vol 6.

Herskowitz I. (1987). Functional inactivation of genes by dominant negative mutations. Nature, 329(6136), 219–222.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., ... & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nature biotechnology, 31(9), 827-832.

Hu, R., Dunn, T. A., Wei, S., Isharwal, S., Veltri, R. W., Humphreys, E., Han, M., Partin, A. W., Vessella, R. L., Isaacs, W. B., Bova, G. S., and Luo, J. (2009). Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify

hormone-refractory prostate cancer. Cancer research, 69(1), 16–22.

Hu, D. G., Hickey, T. E., Irvine, C., Wijayakumara, D. D., Lu, L., Tilley, W. D., Selth, L. A., and Mackenzie, P. I. (2014). Identification of androgen receptor splice variant transcripts in breast cancer cell lines and human tissues. Hormones and cancer, 5(2), 61–71.

Huang, Z., Tamura, M., Sakurai, T., Chuma, S., Saito, T., and Nakatsuji, N. (2000). *In vivo* transfection of testicular germ cells and transgenesis by using the mitochondrially localized jellyfish fluorescent protein gene. FEBS letters, 487(2), 248–251.

Hummel S , Lynn EG, Osanger A, Hirayama S, Nimpf J, Schneider WJ. (2003) Molecular characterization of the first avian LDL receptor: role in sterol metabolism of ovarian follicular cells J Lipid Res. 44(9):1633-42.

Hutchison, J. B., and Steimer, T. (1984). Androgen metabolism in the brain: behavioural correlates. Progress in brain research, 61, 23–51.

Ikawa, M., Tanaka, N., Kao, W. W., and Verma, I. M. (2003). Generation of transgenic mice using lentiviral vectors: a novel preclinical assessment of lentiviral vectors for gene therapy. Molecular therapy : the journal of the American Society of Gene Therapy, 8(4), 666–673.

Immelmann, K. (1962 a). Beiträge zu einer vergleichenden Biologie australischer Prachtfinken. Zool. Jb. Syst. 90: 1–196

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. Bird vocalizations, 61-77.

Itasaki, N., Bel-Vialar, S., and Krumlauf, R. (1999). ‘Shocking’developments in chick embryology: electroporation and *in ovo* gene expression. Nature cell biology, 1(8), E203-E207.

Jackson, A. L., and Linsley, P. S. (2010). Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. Nature reviews Drug discovery, 9(1), 57-67.

Jacobs, E. C., Arnold, A. P., and Campagnoni, A. T. (1999). Developmental regulation of the distribution of aromatase-and estrogen-receptor-mRNA-expressing cells in the zebra finch brain. Developmental neuroscience, 21(6), 453-472.

Jacobs, L. F. (1996). Sexual selection and the brain. Trends in Ecology and Evolution, 11(2), 82-86.

Jaenisch, R., and Mintz, B. (1974). Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA. Proceedings of the national academy of sciences, 71(4), 1250-1254.

Jenster, G., van der Korput, H. A., Trapman, J., and Brinkmann, A. O. (1995). Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. Journal of Biological Chemistry, 270(13), 7341-7346.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. science, 337(6096), 816-821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. elife, 2, e00471.

Jones, F. E., and Stern, D. F. (1999). Expression of dominant-negative ErbB2 in the mammary gland of transgenic mice reveals a role in lobuloalveolar development and lactation. Oncogene, 18(23), 3481-3490.

Joung, J. K., and Sander, J. D. (2013). TALENs: a widely applicable technology for targeted genome editing. Nature reviews Molecular cell biology, 14(1), 49-55.

Jung KM, Kim YM, Keyte AL, Biegler MT, Rengaraj D, Lee HJ, Mello CV, Velho TAF, Fedrigo O, Haase B, Jarvis ED, Han JY (2019) Identification and characterization of primordial germ cells in a vocal learning Neoaves species, the zebra finch. FASEB J 33(12):13825-13836.

Kagami, H., Clark, M. E., Gibbins, A. V., and Etches, R. J. (1995). Sexual differentiation of chimeric chickens containing ZZ and ZW cells in the germline. Molecular reproduction and development, 42(4), 379-387.

Karagenc, L., and Petitte, J. N. (2000). Soluble factors and the emergence of chick primordial germ cells *in vitro*. Poultry science, 79(1), 80-85.

Kay, M. A., Glorioso, J. C., and Naldini, L. (2001). Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nature medicine, 7(1), 33-40.

Kim, Y. H., Perlman, W. R., and Arnold, A. P. (2004). Expression of androgen receptor mRNA in zebra finch song system: developmental regulation by estrogen. Journal of Comparative Neurology, 469(4), 535-547.

Kim, D., Kim, S., Kim, S., Park, J., & Kim, J. S. (2016). Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. Genome research, 26(3), 406-415.

Kino, K., Pain, B., Leibo, S. P., Cochran, M., Clark, M. E., and Etches, R. J. (1997). Production of chicken chimeras from injection of frozen-thawed blastodermal cells. Poultry science, 76(5), 753-760.

Kirn, J. R., and DeVoogd, T. J. (1989). Genesis and death of vocal control neurons during sexual differentiation in the zebra finch. Journal of Neuroscience, 9(9), 3176-3187.

Kistner, A., Gossen, M., Zimmermann, F., Jerecic, J., Ullmer, C., Lübbert, H., and Bujard, H. (1996). Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. Proceedings of the National Academy of Sciences, 93(20), 10933-10938.

Ko, M. C., Van Meir, V., Vellema, M., and Gahr, M. (2020). Characteristics of song, brain-anatomy and blood androgen levels in spontaneously singing female canaries. Hormones and behavior, 117, 104614.

Kochav S, Ginsburg M, Eyal-Giladi H (1980) From cleavage to primitive streak formation: A complementary normal table and a new look at the first stages of the development of the chick: II. Microscopic anatomy and cell population dynamics. *Dev Biol* 79(2):296–308.

Konishi, M., and Akutagawa, E. (1987, January). Hormonal control of cell death in a sexually dimorphic song nucleus in the zebra finch. In Ciba Found. Symp (Vol. 126, pp. 173-85).

Konishi, M., and Gurney, M. E. (1982). Sexual differentiation of brain and behaviour. Trends in neurosciences, 5, 20-23.

Konishi, M. (1965). The Role of Auditory Feedback in the Control of Vocalization in the White-Crowned Sparrow 1. Zeitschrift für Tierpsychologie, 22(7), 770-783.

Konopka G, Roberts TF (2016) Insights into the Neural and Genetic Basis of Vocal

Communication. *Cell* 164(6):1269–1276.

Kousteni, S., Bellido, T., Plotkin, L. I., O'brien, C. A., Bodenner, D. L., Han, L., ... and Manolagas, S. C. (2001). Nongenotropic, sex-nonspecific signaling through the estrogen or androgen receptors: dissociation from transcriptional activity. Cell, 104(5), 719-730.

Kurz, E. M., Sengelaub, D. R., and Arnold, A. P. (1986). Androgens regulate the dendritic length of mammalian motoneurons in adulthood. Science, 232(4748), 395-398.

Lavial, F., Acloque, H., Bertocchini, F., MacLeod, D. J., Boast, S., Bachelard, E., ... and Pain, B. (2007). The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. Development, 134(19), 3549-3563.

Lavitrano, M., Camaioni, A., Fazio, V. M., Dolci, S., Farace, M. G., and Spadafora, C. (1989). Sperm cells as vectors for introducing foreign DNA into eggs: genetic transformation of mice. Cell, 57(5), 717-723.

Lee, J., Kim, D. H., and Lee, K. (2020). Current Approaches and Applications in Avian Genome Editing. International Journal of Molecular Sciences, 21(11), 3937.

Leranth, C., Petnehazy, O., and MacLusky, N. J. (2003). Gonadal hormones affect spine synaptic density in the CA1 hippocampal subfield of male rats. Journal of Neuroscience, 23(5), 1588-1592.

Liao Y, Smyth G, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, *Bioinformatics* 30 (7): 923–930 DOI:

Liu WC, Kohn J, Szwed SK, Pariser E, Sepe S, Haripal B, Oshimori N, Marsala M, Miyanohara A, Lee R (2015) Human mutant huntingtin disrupts vocal learning in transgenic songbirds. *Nat Neurosci* 18:1617-1622.

Lois, C., Hong, E. J., Pease, S., Brown, E. J., and Baltimore, D. (2002). Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. Science, 295(5556), 868-872.

Love M.I, Huber W, Anders, S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550

Lovell, P. V., Huizinga, N. A., Friedrich, S. R., Wirthlin, M., and Mello, C. V. (2018). The constitutive differential transcriptome of a brain circuit for vocal learning. BMC

genomics, 19(1), 1-27.

Lu, C., and Luo, J. (2013). Decoding the androgen receptor splice variants. Translational andrology and urology, 2(3), 178.

Lu, C., Brown, L. C., Antonarakis, E. S., Armstrong, A. J., and Luo, J. (2020). Androgen receptor variant-driven prostate cancer II: advances in laboratory investigations. Prostate cancer and prostatic diseases, 23(3), 381-397.

Luo, M., Yu, Y., Kim, H., Kudrna, D., Itoh, Y., Agate, R. J., ... and Arnold, A. P. (2006). Utilization of a zebra finch BAC library to determine the structure of an avian androgen receptor genomic region. Genomics, 87(1), 181-190.

Macdonald J, Glover JD, Taylor L, Sang HM, McGrew MJ (2010) Characterisation and Germline Transmission of Cultured Avian Primordial Germ Cells. *PLoS One* 5(11):e15518.

Madisen, L., Mao, T., Koch, H., Zhuo, J. M., Berenyi, A., Fujisawa, S., ... and Zeng, H. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nature neuroscience, 15(5), 793-802.

Maione, B., Lavitrano, M., Spadafora, C., and Kiessling, A. A. (1998). Sperm-mediated gene transfer in mice. Molecular Reproduction and Development: Incorporating Gamete Research, 50(4), 406-409.

Mak S, Alev C, Nagai H, Wrabel A, Matsuoka Y, Honda A, Sheng G, Ladher RK (2015) Characterization of the finch embryo supports evolutionary conservation of the naive stage of development in amniotes. eLife 4:e07178.

Makoolati, Z., Movahedin, M., and Forouzandeh-Moghadam, M. (2016). Proliferation in culture of primordial germ cells derived from embryonic stem cell: induction by retinoic acid. Bioscience reports, 36(6).

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., ... and Church, G. M. (2013). RNA-guided human genome engineering via Cas9. Science, 339(6121), 823-826.

Manova, K., Paynton, B. V., and Bachvarova, R. F. (1992). Expression of activins and TGFβ1 and β2 RNAs in early postimplantation mouse embryos and uterine decidua. Mechanisms of development, 36(3), 141-152.

Margoliash, D. (1997). Functional organization of forebrain pathways for song production and perception. Journal of neurobiology, 33(5), 671-693.

Marler, P. (1970). Birdsong and speech development: Could there be parallels? There may be basic rules governing vocal learning to which many species conform, including man. American scientist, 58(6), 669-673.

Matias, P. M., Donner, P., Coelho, R., Thomaz, M., Peixoto, C., Macedo, S., ... and Carrondo, M. A. (2000). Structural evidence for ligand specificity in the binding domain of the human androgen receptor: implications for pathogenic gene mutations. Journal of Biological Chemistry, 275(34), 26164-26171.

Matsui, Y., Toksoz, D., Nishikawa, S., Nishikawa, S. I., Williams, D., Zsebo, K., and Hogan, B. L. (1991). Effect of Steel factor and leukaemia inhibitory factor on murine primordial germ cells in culture. Nature, 353(6346), 750-752.

Matsumoto, T., and Claesson-Welsh, L. (2001). VEGF receptor signal transduction. Science's STKE : signal transduction knowledge environment, 2001(112), re21.

McCarthy MM, Arnold AP (2011) Reframing sexual differentiation of the brain. *Nat Neurosci* 14(6):677–683.

McCarthy, M. M. (2016). Multifaceted origins of sex differences in the brain. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1688), 20150106.

McDonald, K. S., and Kirn, J. R. (2012). Anatomical plasticity in the adult zebra finch song system. Journal of Comparative Neurology, 520(16), 3673-3686.

McGrew, M. J., Sherman, A., Ellard, F. M., Lillico, S. G., Gilhooley, H. J., Kingsman, A. J., ... and Sang, H. (2004). Efficient production of germline transgenic chickens using lentiviral vectors. EMBO reports, 5(7), 728-733.

McLellan, M. A., Rosenthal, N. A., and Pinto, A. R. (2017). Cre-loxP-mediated recombination: General principles and experimental considerations. Current protocols in mouse biology, 7(1), 1-12.

Mello CV (2014) The zebra finch, Taeniopygia guttata: an avian model for investigating the neurobiological basis of vocal learning. *Cold Spring Harb Protoc* 2014(12):1237–1242.

Ménoret S, Tesson L, Remy S, Usal C, Ouisse LH, Brusselle L, Chenouard V, Anegon I (2017) Advances in transgenic animal models and techniques. *Transgenic Res* 26, 703–708.

Metzdorf, R., Gahr, M., and Fusani, L. (1999). Distribution of aromatase, estrogen receptor, and androgen receptor mRNA in the forebrain of songbirds and nonsongbirds. Journal of Comparative Neurology, 407(1), 115-129.

Meyer, M., de Angelis, M. H., Wurst, W., and Kühn, R. (2010). Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proceedings of the National Academy of Sciences, 107(34), 15022-15026.

Michalkiewicz, M., Michalkiewicz, T., Geurts, A. M., Roman, R. J., Slocum, G. R., Singer, O., ... and Cowley Jr, A. W. (2007). Efficient transgenic rat production by a lentiviral vector. American Journal of Physiology-Heart and Circulatory Physiology, 293(1), H881-H894.

Mizuarai, S., Ono, K. I., Yamaguchi, K., Nishijima, K. I., Kamihira, M., and Iijima, S. (2001). Production of transgenic quails with high frequency of germ-line transmission using VSV-G pseudotyped retroviral vector. Biochemical and biophysical research communications, 286(3), 456-463.

Mooney, R., and Rao, M. (1994). Waiting periods versus early innervation: the development of axonal connections in the zebra finch song system. Journal of Neuroscience, 14(11), 6532-6543.

Mooney, R. (1999). Sensitive periods and circuits for learned birdsong. Current opinion in neurobiology, 9(1), 121-127.

Mooney (2020) The neurobiology of innate and learned vocalizations in rodents and songbirds. *Curr Opin Neurobiol* 2020 (64) 24-31.

Morris, D. (1954). The reproductive behaviour of the zebra finch (Poephila guttata), with special reference to pseudofemale behaviour and displacement activities. Behaviour, 6(1), 271-322.

Mozdziak, P. E., Borwornpinyo, S., McCoy, D. W., and Petitte, J. N. (2003). Development of transgenic chickens expressing bacterial β-galactosidase. Developmental dynamics: an official publication of the American Association of Anatomists, 226(3), 439-445.

Murray JR, Varian-Ramos CW, Welch ZS, Saha MS. (2013) Embryological staging of the Zebra Finch, Taeniopygia guttata. *J Morphol*. 274(10):1090-110.

Nakamura M, Kuwana T, Miyayama Y, Fujimoto T (1988) Extragonadal distribution of primordial germ cells in the early chick embryo. *Anat Rec* 222(1):90–94.

Nakamura, Y., Yamamoto, Y., Usui, F., Atsumi, Y., Ito, Y., Ono, T., ... and Tagami, T. (2008). Increased proportion of donor primordial germ cells in chimeric gonads by sterilisation of recipient embryos using busulfan sustained-release emulsion in chickens. Reproduction, Fertility and Development, 20(8), 900-907.

Nalbandian, A., Dettin, L., Dym, M., and Ravindranath, N. (2003). Expression of vascular endothelial growth factor receptors during male germ cell differentiation in the mouse. Biology of Reproduction, 69(3), 985-994.

Nealen, P. M. (2005). An interspecific comparison using immunofluorescence reveals that synapse density in the avian song system is related to sex but not to male song repertoire size. Brain research, 1032(1-2), 50-62.

Nieder, A., and Mooney, R. (2020). The neurobiology of innate, volitional and learned vocalizations in mammals and birds. Philosophical Transactions of the Royal Society B, 375(1789), 20190054.

Nikolic J, Belot L, Raux H, Legrand P, Gaudin Y, Albertini A. (2018) Structural basis for the recognition of LDL-receptor family members by VSV glycoprotein. Nat Commun 9, 1029.

Nishii K, Nakaseko C, Jiang M, Shimizu N, Takeuchi M, Schneider W, Bujo H. (2013) The soluble form of LR11 protein is a regulator of hypoxia-induced, urokinase-type plasminogen activator receptor (uPAR)-mediated adhesion of immature hematological cells. *J Biol Chem*. 288(17):11877-11886.

Nixdorf-Bergweiler, B. E. (1996). Divergent and parallel development in volume sizes of telencephalic song nuclei in and female zebra finches. Journal of Comparative Neurology, 375(3), 445-456.

Nordeen, E. J., and Nordeen, K. W. (1989). Estrogen stimulates the incorporation of new neurons into avian song nuclei during adolescence. Developmental Brain Research, 49(1), 27-32.

Nordeen, K. W., Nordeen, E. J., and Arnold, A. P. (1987). Estrogen accumulation in zebra finch song control nuclei: implications for sexual differentiation and adult activation of song behavior. Journal of neurobiology, 18(6), 569-582.

Norton, P., Barschke, P., Scharff, C., and Mendoza, E. (2019). Differential song deficits after lentivirus-mediated knockdown of FoxP1, FoxP2, or FoxP4 in area X of juvenile zebra finches. Journal of Neuroscience, 39(49), 9782-9796.

Nottebohm, F., and Arnold, A. P. (1976). Sexual dimorphism in vocal control areas of the songbird brain. Science, 194(4261), 211-213.

Nottebohm, F., Stokes, T. M., and Leonard, C. M. (1976). Central control of song in the canary, Serinus canarius. Journal of Comparative Neurology, 165(4), 457-486.

Nottebohm, F., Nottebohm, M. E., Crane, L. A., and Wingfield, J. C. (1987). Seasonal changes in gonadal hormone levels of adult male canaries and their relation to song. Behavioral and neural biology, 47(2), 197-211.

Nottebohm, F., Paton, J. A., and Kelley, D. B. (1982). Connections of vocal control nuclei in the canary telencephalon. Journal of Comparative Neurology, 207(4), 344-357.

Oishi I, Yoshii K, Miyahara D, Kagami H, Tagami T (2016) Targeted mutagenesis in chicken using CRISPR/Cas9 system. *Sci Rep* 6(1):23980.

Oishi I, Yoshii K, Miyahara D, Tagami T (2018) Efficient production of human interferon beta in the white of eggs from ovalbumin gene-targeted hens. *Sci Rep* 8:10203.

Paton JA, Nottebohm FN (1984) Neurons generated in the adult brain are recruited into functional circuits. *Science* 225(4666):1046–1048.

Pauklin, S., and Vallier, L. (2015). Activin/Nodal signalling in stem cells. Development, 142(4), 607-619.

Pesce, M., Farrace, M. G., Piacentini, M., Dolci, S., and De Felici, M. (1993). Stem cell factor and leukemia inhibitory factor promote primordial germ cell survival by suppressing programmed cell death (apoptosis). Development, 118(4), 1089-1094.

Peterson, R. S., Yarram, L., Schlinger, B. A., and Saldanha, C. J. (2005). Aromatase is pre-synaptic and sexually dimorphic in the adult zebra finch brain. Proceedings of the Royal Society B: Biological Sciences, 272(1576), 2089-2096.

Petitte, J. N., Clark, M. E., Liu, G., Gibbins, A. V., and Etches, R. J. (1990). Production of somatic and germline chimeras in the chicken by transfer of early blastodermal cells. Development, 108(1), 185-189.

Pfeifer, A. (2004). Lentiviral transgenesis. Transgenic research, 13(6), 513-522.

Pfeifer, A., Ikawa, M., Dayn, Y., and Verma, I. M. (2002). Transgenesis by lentiviral vectors: lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. Proceedings of the National Academy of Sciences, 99(4), 2140-2145.

Podmokła, E., Drobniak, S. M., & Rutkowska, J. (2018). Chicken or egg? Outcomes of experimental manipulations of maternally transmitted hormones depend on administration method–a meta-analysis. Biological Reviews, 93(3), 1499-1517.

Prather JF, Okanoya K, Bolhuis JJ (2017) Brains for birds and babies: Neural parallels between birdsong and speech acquisition. *Neurosci Biobehav Rev* 81:225–237.

Pytte, C. L., Gerson, M., Miller, J., and Kirn, J. R. (2007). Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. Developmental neurobiology, 67(13), 1699-1720.

Quispe, R., Sèbe, F., da Silva, M. L., and Gahr, M. (2016). Dawn-song onset coincides with increased HVC androgen receptor expression but is decoupled from high circulating testosterone in an equatorial songbird. Physiology and behavior, 156, 1-7.

Rai, N. K., Tripathi, K., Sharma, D., and Shukla, V. K. (2005). Apoptosis: a basic physiologic process in wound healing. The international journal of lower extremity wounds, 4(3), 138-144.

Rao, P., and Monks, D. A. (2009). A tetracycline-inducible and skeletal muscle-specific Cre recombinase transgenic mouse. Developmental neurobiology, 69(6), 401-406.

Raucci F, Fuet A, Pain B (2015) *In vitro* generation and characterization of chicken long-term germ cells from different embryonic origins. *Theriogenology* 84(5):732-742

Reiner, A., Perkel, D. J., Bruce, L. L., Butler, A. B., Csillag, A., Kuenzel, W., ... and Jarvis, E. D. (2004). Revised nomenclature for avian telencephalon and some related brainstem nuclei. Journal of Comparative Neurology, 473(3), 377-414.

Resnick, J. L., Bixler, L. S., Cheng, L., and Donovan, P. J. (1992). Long-term proliferation

of mouse primordial germ cells in culture. Nature, 359(6395), 550-551.

Riebel, K., Odom, K. J., Langmore, N. E., and Hall, M. L. (2019). New insights from female bird song: towards an integrated approach to studying male and female communication roles. Biology letters, 15(4), 20190059.

Roper, A., and Zann, R. (2006). The onset of song learning and song tutor selection in fledgling zebra finches. Ethology, 112(5), 458-470.

Salter, D. W., Smith, E. J., Hughes, S. H., Wright, S. E., and Crittenden, L. B. (1987). Transgenic chickens: insertion of retroviral genes into the chicken germ line. Virology, 157(1), 236-240.

Sander, J. D., and Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nature biotechnology, 32(4), 347-355.

Santa-Gonzalez, G. A., Gomez-Molina, A., Arcos-Burgos, M., Meyer, J. N., and Camargo, M. (2016). Distinctive adaptive response to repeated exposure to hydrogen peroxide associated with upregulation of DNA repair genes and cell cycle arrest. Redox biology, 9, 124-133.

Sato, T., Matsumoto, T., Kawano, H., Watanabe, T., Uematsu, Y., Sekine, K., ... & Kato, S. (2004). Brain masculinization requires androgen receptor function. Proceedings of the National Academy of Sciences, 101(6), 1673-1678.

Scaltriti, M., and Baselga, J. (2006). The epidermal growth factor receptor pathway: a model for targeted therapy. Clinical cancer research, 12(18), 5268-5272.

Schielzeth, H., and Bolund, E. (2010). Patterns of conspecific brood parasitism in zebra finches. Animal behaviour, 79(6), 1329-1337.

Schreihofer, D. A., Duong, P., and Cunningham, R. L. (2018). N-terminal truncations in sex steroid receptors and rapid steroid actions. Steroids, 133, 15-20.

Schusser B, Collarini EJ, Pedersen D, Yi H, Ching K, Izquierdo S, Thoma T, Lettmann S, Kaspers B, Etches RJ, van de Lavoir MC, Harriman W, Leighton PA (2016) Expression of heavy chain-only antibodies can support B-cell development in light chain knockout chickens. *Eur J Immunol* 46(9):2137–2148.

Schusser B, Collarini EJ, Yi H, Izquierdo SM, Fesler J, Pedersen D, Klasing KC, Kaspers

B, Harriman WD, van de Lavoir MC, Etches RJ, Leighton PA. (2013) Immunoglobulin knockout chickens via efficient homologous recombination in primordial germ cells. *Proc Natl Acad Sci U S A* 110(50):20170–20175.

Seibler, J., Kleinridders, A., Küter-Luks, B., Niehaves, S., Brüning, J. C., and Schwenk, F. (2007). Reversible gene knockdown in mice using a tight, inducible shRNA expression system. Nucleic acids research, 35(7), e54.

Shaffer, P. L., Jivan, A., Dollins, D. E., Claessens, F., and Gewirth, D. T. (2004). Structural basis of androgen receptor binding to selective androgen response elements. Proceedings of the National Academy of Sciences of the United States of America, 101(14), 4758–4763.

Sid H, Schusser B (2018) Applications of Gene Editing in Chickens: A New Era Is on the Horizon. *Front Genet* 9:456.

Slagsvold, T., Kraus, I., Bentzen, T., Palvimo, J., and Saatcioglu, F. (2000). Mutational analysis of the androgen receptor AF-2 (activation function 2) core domain reveals functional and mechanistic differences of conserved residues compared with other nuclear receptors. Molecular Endocrinology, 14(10), 1603-1617.

Slater, P. J., Eales, L. A., and Clayton, N. S. (1988). Song learning in zebra finches (Taeniopygia guttata): progress and prospects. In Advances in the Study of Behavior (Vol. 18, pp. 1-34). Academic Press.

Smih, F., Rouet, P., Romanienko, P. J., and Jasin, M. (1995). Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic acids research, 23(24), 5012-5019.

Smith, K., and Spadafora, C. (2005). Sperm-mediated gene transfer: Applications and implications. Bioessays, 27(5), 551-562.

Spritzer, M. D., and Galea, L. A. (2007). Testosterone and dihydrotestosterone, but not estradiol, enhance survival of new hippocampal neurons in adult male rats. Developmental neurobiology, 67(10), 1321-1333.

Sun, S., Sprenger, C. C., Vessella, R. L., Haugk, K., Soriano, K., Mostaghel, E. A., Page, S. T., Coleman, I. M., Nguyen, H. M., Sun, H., Nelson, P. S., and Plymate, S. R. (2010). Castration resistance in human prostate cancer is conferred by a frequently occurring

androgen receptor splice variant. The Journal of clinical investigation, 120(8), 2715–2730.

Sung, Y. H., Baek, I. J., Kim, D. H., Jeon, J., Lee, J., Lee, K., ... and Lee, H. W. (2013). Knockout mice created by TALEN-mediated gene targeting. Nature biotechnology, 31(1), 23-24.

Tan, M. H., Li, J., Xu, H. E., Melcher, K., and Yong, E. L. (2015). Androgen receptor: structure, role in prostate cancer and drug discovery. Acta pharmacologica Sinica, 36(1), 3–23.

Tararova, N. D., Narizhneva, N., Krivokrisenko, V., Gudkov, A. V., and Gurova, K. V. (2007). Prostate cancer cells tolerate a narrow range of androgen receptor expression and activity. The Prostate, 67(16), 1801-1815.

Taylor L, Carlson DF, Nandi S, Sherman A, Fahrenkrug SC, McGrew MJ. (2017) Efficient TALEN-mediated gene targeting of chicken primordial germ cells. *Development* 144(5):928 – 934.

Tesson, L., Usal, C., Ménoret, S., Leung, E., Niles, B. J., Remy, S., ... and Cost, G. J. (2011). Knockout rats generated by embryo microinjection of TALENs. Nature biotechnology, 29(8), 695-696.

Thomas, K. R., and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell, 51(3), 503-512.

Thompson, J. A., Wu, W., Bertram, R., and Johnson, F. (2007). Auditory-dependent vocal recovery in adult male zebra finches is facilitated by lesion of a forebrain pathway that includes the basal ganglia. The Journal of neuroscience : the official journal of the Society for Neuroscience, 27(45), 12308–12320.

Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L., and Melton, D. W. (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell, 56(2), 313-321.

Thorpe, W. H. (1954). The process of song-learning in the chaffinch as studied by means of the sound spectrograph. Nature, 173(4402), 465-469.

Tian, R., Yang, S., Zhu, Y., Zou, S., Li, P., Wang, J., ... and Li, Z. (2016). VEGF/VEGFR2 signaling regulates germ cell proliferation *in vitro* and promotes mouse testicular

regeneration *in vivo*. Cells Tissues Organs, 201(1), 1-13.

Tiscornia, G., Singer, O., Ikawa, M., and Verma, I. M. (2003). A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. Proceedings of the National Academy of Sciences, 100(4), 1844-1848.

Tramontin, A. D., and Brenowitz, E. A. (2000). Seasonal plasticity in the adult brain. Trends in neurosciences, 23(6), 251-258.

Tramontin, A. D., Wingfield, J. C., and Brenowitz, E. A. (2003). Androgens and estrogens induce seasonal-like growth of song nuclei in the adult songbird brain. Journal of neurobiology, 57(2), 130-140.

Trefil, P., Micáková, A., Mucksová, J., Hejnar, J., Poplstein, M., Bakst, M. R., ... and Brillard, J. P. (2006). Restoration of spermatogenesis and male fertility by transplantation of dispersed testicular cells in the chicken. Biology of Reproduction, 75(4), 575-581.

Trefil, P., Aumann, D., Koslová, A., Mucksová, J., Benešová, B., Kalina, J., ... and Hejnar, J. (2017). Male fertility restored by transplanting primordial germ cells into testes: A new way towards efficient transgenesis in chicken. Scientific reports, 7(1), 1-9.

Tschirren, B. (2015). Differential effects of maternal yolk androgens on male and female offspring: a role for sex-specific selection?. PLoS One, 10(7), e0133673.

Tyack, S. G., Jenkins, K. A., O'Neil, T. E., Wise, T. G., Morris, K. R., Bruce, M. P., ... and Doran, T. J. (2013). A new method for producing transgenic birds via direct *in vivo* transfection of primordial germ cells. Transgenic Research, 22(6), 1257-1264.

Urven LE, Erickson CA, Abbott UK, McCarrey JR (1988) Analysis of germ line development in the chick embryo using an anti-mouse EC cell antibody. *Development* 103(2):299–304.

Van De Lavoir MC, Diamond JH, Leighton PA, Mather-Love C, Heyer BS, Bradshaw R, Kerchner A, Hooi LT, Gessaro TM, Swanberg SE, Delany ME, Etches RJ. (2006) Germline transmission of genetically modified primordial germ cells. *Nature* 441:766-769.

Vieillot, L.J.P. (1817). Nouveau Dictionnaire d'Histoire naturelle, appliquée aux arts, à l'agriculture, à l'économie rurale et domestique, à la médecine. Tomo 8. 610 pp. 10 tt. París: Deterville.

Voigt, C., and Gahr, M. (2011). Social status affects the degree of sex difference in the songbird brain. PloS one, 6(6), e20723.

Voigt, C., and Leitner, S. (2008). Seasonality in song behaviour revisited: seasonal and annual variants and invariants in the song of the domesticated canary (Serinus canaria). Hormones and behavior, 54(3), 373-378.

Wach, S., Taubert, H., and Cronauer, M. (2020). Role of androgen receptor splice variants, their clinical relevance and treatment options. World journal of urology, 38(3), 647–656.

Wade, J., and Arnold, A. P. (2004). Sexual differentiation of the zebra finch song system. Annals of the New York Academy of Sciences, 1016(1), 540-559.

Wade, J., Schlinger, B. A., Hodges, L., and Arnold, A. P. (1994). Fadrozole: a potent and specific inhibitor of aromatase in the zebra finch brain. General and comparative endocrinology, 94(1), 53-61.

Wadosky, K. M., and Koochekpour, S. (2017). Androgen receptor splice variants and prostate cancer: From bench to bedside. Oncotarget, 8(11), 18550.

Wang, R. N., Green, J., Wang, Z., Deng, Y., Qiao, M., Peabody, M., ... and Shi, L. L. (2014). Bone Morphogenetic Protein (BMP) signaling in development and human diseases. Genes and diseases, 1(1), 87-105.

Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Künstner, A., ... and Wilson, R. K. (2010). The genome of a songbird. Nature, 464(7289), 757-762.

Watson, P. A., Chen, Y. F., Balbas, M. D., Wongvipat, J., Socci, N. D., Viale, A., ... and Sawyers, C. L. (2010). Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. Proceedings of the national academy of sciences, 107(39), 16759-16765.

Weiss, B., Faus, H., and Haendler, B. (2007). Phylogenetic conservation of the androgen receptor AR45 variant form in placental mammals. Gene, 399(2), 105-111.

Veres, A., Gosis, B. S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., ... & Musunuru, K. (2014). Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell stem cell, 15(1), 27-30.

Whyte J, Glover JD, Woodcock M, Brzeszczynska J, Taylor L, Sherman A, Kaiser P, McGrew MJ (2015) FGF, Insulin, and SMAD Signaling Cooperate for Avian Primordial Germ Cell Self-Renewal. *Stem Cell Rep* 5(6):1171–1182.

Whitelaw, C. B. A. (2004). Transgenic livestock made easy. TRENDS in Biotechnology, 22(4), 157-159.

Wild, J. M. (2004) Functional neuroanatomy of the sensorimotor control of singing. Ann N Y Acad Sci. 1016, 438-62.

Xu, D., Zhan, Y., Qi, Y., Cao, B., Bai, S., Xu, W., Gambhir, S. S., Lee, P., Sartor, O., Flemington, E. K., Zhang, H., Hu, C. D., & Dong, Y. (2015). Androgen Receptor Splice Variants Dimerize to Transactivate Target Genes. Cancer research, 75(17), 3663–3671.

Yamamura, T., Barker, J. M., Balthazart, J., and Ball, G. F. (2011). Androgens and estrogens synergistically regulate the expression of doublecortin and enhance neuronal recruitment in the song system of adult female canaries. Journal of Neuroscience, 31(26), 9649-9657.

Yeh, S., Tsai, M. Y., Xu, Q., Mu, X. M., Lardy, H., Huang, K. E., Lin, H., Yeh, S. D., Altuwaijri, S., Zhou, X., Xing, L., Boyce, B. F., Hung, M. C., Zhang, S., Gan, L., & Chang, C. (2002). Generation and characterization of androgen receptor knockout (ARKO) mice: an in vivo model for the study of androgen functions in selective tissues. Proceedings of the National Academy of Sciences of the United States of America, 99(21), 13498–13503.

Zann, R., and Cash, E. (2008). Developmental stress impairs song complexity but not learning accuracy in non-domesticated zebra finches (Taeniopygia guttata). Behavioral Ecology and Sociobiology, 62(3), 391-400.

Zhang, Z., Sun, P., Yu, F., Yan, L., Yuan, F., Zhang, W., Wang, T., Wan, Z., Shao, Q., and Li, Z. (2012). Transgenic quail production by microinjection of lentiviral vector into the early embryo blood vessels. PloS one, 7(12), e50817.

Acknowledgements

I would like to first thank my PhD supervisors and PAC members Manfred Gahr, Falk Dittrich, Moritz Hertel and Mike McGrew. To Manfred, thank you for giving me the opportunity to dive into the academic world doing research on a topic I always dreamed of working on. Thank you for the inspirational ideas and comments, for the guidance but also the freedom to explore different approaches and perspectives to find the right story to tell. To Falk, thank you for being an amazing mentor, for helping me make sense of the data, for all the insightful discussions throughout the years, for finding clever solutions to the problems encountered along the way, thank you for always being available to answer to all my questions, even the silliest ones, thank you for the enormous patience. To Moritz, thank you for introducing me to the world of science, for teaching me most of the lab techniques I know today with patience and kindness, thank you for the support along the years and for your positive encouraging support during the ups and specially the downs of this PhD journey. To Mike McGrew thank you for being part of my PAC, for taking the time to travel to Seewiesen for our meetings, for all the insightful comments. Thank you for sharing your knowledge with me, for teaching me all the techniques used to develop transgenic chickens and for helping me adapt them to songbirds. Also a special thank you to all your team at the Roslin institute for receiving me with opened arms and specially Lorna Taylor and Sunil Nandi for patiently teaching me a lot of protocols and useful tips.

Thank you to all my collaborators that made possible a lot of the experiments done in this thesis. Thank you to Staffan Hildebrand and Alexander Pfeifer for developing the viruses used in this thesis. Thank you to Bernd Timmermann for performing the next generation sequencing of our samples and to Carolina Frankl-Vilches for helping me understand how to analyse sequencing data.

I would also like to thank my research school, the IMPRS for organismal biology, for offering a space to develop new skills and to meet people going through the same PhD journey with whom we became amazing friend. Thank you for the support to visit conferences and attend to courses and symposiums. Special thanks to Mäggi Hieber-Ruiz, Corinna Loës and Francisca Rosa Mende our IMPRS coordinators for all the help with bureaucracy and organization.

Thank you to the whole Gahr lab, specially to the technicians Christina Reusch, Anja Lohrentz, Judith Kammerlander, Antje Bakker, for all the support, the patience and the wonderful working environment that made the PhD process a lot more enjoyable.

Thank you to the animal care takers David Witkowski, Frances Weigl and Frank Lehmann for their support caring for the birds with kindness and patience.

A big thank you to all the great friends I made at the IMPRS with whom I shared experiences and advices. A special thank you to Amanda M., Lucía M., Luisana C. and Mariana D. for the amazing friendship, the enormous emotional support, the great memories and all the girls nights with the most hilarious conversations. Thank you to Julia C. for always being there for me, for all the coffee breaks, the emotional support and life changing conversations. Thank you to all the amazing friends I made in Seewiesen, Daniel M., Esteban B., Glenn C., Gianina U., Juliana A., Kamila L., Keren S., Laurie O., Lina G., Maggie K., Maria M., Mekhla R., Nicolas A., Qiaoyi L., Sandra E., Theresa H. and Yifan P., for all their support and great memories, and for making the PhD experience a wonderful one.

Gracias a mis amigos de la vida Aldi, Andre, Fausto, Hugo, Ine, Juan, Lu, Marian y Mily por el apoyo incondicional, por estar siempre presentes en mi vida, estando lejos o cerca, durante los momentos lindos y los mas difíciles.

Gracias a mis amigos del telepuerto, Andre, Andre, Cristian, Enzi, Ger, Juan, Lean, Pablo, Tami y Vero, por ser ese rincón de argentinidad que nos mantiene cerca de nuestras raíces, por los mates y las cenas y los recuerdos de esta experiencia alemana tan genial.

Gracias a Tomas por los últimos 12 años, por ser un compañero de aventuras excelente, por aguantarme en los momentos difíciles y por bajarme a la realidad en los momentos mas estresantes del doctorado recordándome que el trabajo no es lo mas importante.

Gracias a mi familia mamá, papá y Diane, por ser incondicionales, por acompañarme siempre, por estar lejos pero a la vez muy cerca, por dejarme volar y seguir mis sueños y por haberme ayudado a ser la persona que soy hoy. Son maravillosos, los amo!

Yuki mi compañera de aventuras de cuatro patas, me hace muy feliz que estés en mi vida y siempre voy a estar agradecida por tu compañía y amor infinito de perro.

Author Contributions

Chapter I

I.G., F.D., M.H., M.M., M.G. conceived the study, M.G.,M.M. and A.P provided instruments, materials and reagents. I.G. performed the experiments. S.H. produced the lentiviral vectors. CFV contributed to the analysis of the sequencing data. I.G. wrote the manuscript and all coauthors contributed to manuscript revision. I.G. is a member of the International Max Planck Research School (IMPRS) for Organismal Biology and the work was funded by the Max Planck Society.

Chapter II

I.G., F.D., M.H., M.G. conceived the study, M.G. provided instruments, materials and reagents. I.G. performed the experiments. CFV contributed to the analysis of the sequencing data. I.G. wrote the manuscript and all coauthors contributed to manuscript revision. I.G. is a member of the International Max Planck Research School (IMPRS) for Organismal Biology and the work was funded by the Max Planck Society.

Chapter III

I.G., F.D., M.H., M.G. conceived the study, M.G. provided instruments, materials and reagents. I.G., F.D., M.H. contributed to experiments and interpretation of results. I.G. generated all figures, wrote the manuscript and all coauthors contributed to manuscript revision. I.G. is a member of the International Max Planck Research School (IMPRS) for Organismal Biology and the work was funded by the Max Planck Society.