Genomic resources for the flatworm model organism *Macrostomum lignano*

Daniil Simanov

ISBN: 978-94-6182-425-7

Layout and printing: Off Page

Cover: Two Macrostomum worms and DNA fragment; by Dan and Ira.

Genomic resources for the flatworm model organism *Macrostomum lignano*

Genomische middelen voor de platworm modelorganisme Macrostomum lignano

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 15 april 2014 des middags te 12:45 uur

door

Daniil Simanov

geboren op 11 mei 1986

te Moskou, Rusland

Promotoren: Prof. dr. E.P.J.G Cuppen
Prof. dr. E. Berezikov

The research described in this thesis was performed at the Hubrecht Institute of The Royal Academy of Arts and Sciences (KNAW) within the framework of the Graduate School of Cancer, Stem Cells and Developmental Biology (CSnD) in Utrecht.

To my colleagues -
all the brave, dedicated and patient people I was lucky to work side by side with.

CONTENTS

CHAPTER 1

INTRODUCTION

PREFACE

The term "stem cell" was first proposed for scientific use in 1908 by the Russian histologist Aleksander Maksimov. Since then we have learned a lot, and advances in our knowledge and understanding of stem cells have changed the way we see human biology, physiology, deceases and ways to treat them. Research into the role of stem cells in mammalian organisms has expanded rapidly over the past 25 years. Martin Evans, Mario Capecchi and Oliver Smithies in 2007 and Shinya Yamanaka with John Gurdon in 2012 were recognized for their contribution to stem-cell research by being award Nobel Prizes for Medicine and Physiology.

Today we know a lot about role of stem cells in embryonic and post-embryonic development, regeneration, cancer progression, ageing, and some disorders. And for the last 15 years there are a lot of talks in the scientific community and outside of it about stem cell therapy and its potential to dramatically change treatment of a great number of human deceases. Yet, as of 2013, the only established therapeutic use of stem cells is bone marrow transplantation – technique that was first successfully performed already in 1968.

A lot of challenges and questions still remain in the field, and a lot of work is still to be done before stem cell therapy becomes real and can reach its full potential. Developing a new attractive model to study adult stem cells *in vivo* is but a short step towards this goal, and that is what this thesis is about.

1. STEM CELL

Stem cells are undifferentiated cells that share, regardless of the type, two defining properties. First, they are capable of renewing themselves through cell division, sometimes after long periods of inactivity. Second, under certain physiologic or experimental conditions, they can be induced to become tissue- or organ-specific cells with special functions (Macarthur et al., 2009). This unique character makes them crucial for embryonic and post-embryonic development, as well as for the tissue homeostasis throughout life (compensation of cell mass lost through apoptosis or necrosis) and regeneration of damaged tissues. Stem cells can divide both symmetrically and asymmetrically (Morrison and Kimble, 2006). During early embryonic development or recovery after injury, when cells are rapidly expanding in number, stem cell division stays symmetrical i.e. each cell divides to give rise to daughter cells each with virtually same potential. Later in development or regeneration process cells divide asymmetrically with one of the two newly formed cells keeping its stemness, whereas the fate of the other one is to differentiate terminally (Neumüller and Knoblich, 2009).

At every moment throughout the life of multicellular organism the number of stem cells, their position in the tissues, proliferation activity and differentiation process of their descendants must remain under tight control. Malfunction of stem cells can lead either to tumor formation (Houghton et al., 2007; Risitano et al., 2007; DeGregori, 2011) or to premature ageing (Hartwig and Collares, 2013; Fulle et al., 2012; Beltrami et al., 2011; Sharpless and DePinho, 2007). Proper functioning of stem cells depends on both intrinsic (cell autonomous) and extrinsic (cellular contacts, secreted substances and other environmental signals) regulatory mechanisms (Walker et al., 2009), so not only the stem cell, but also its specialized micro-

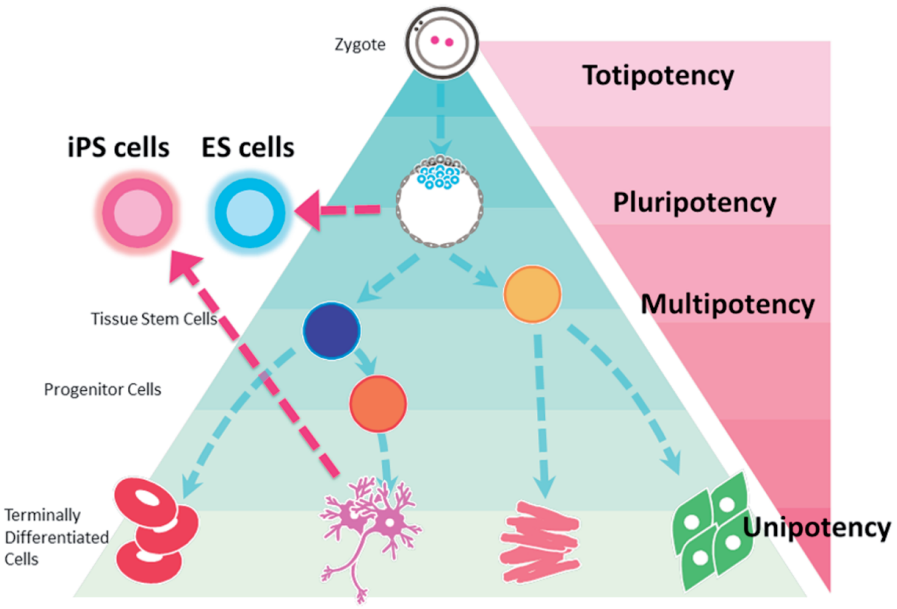


Figure 1. Hierarchical potential of stem cell development. Zygote, the only true totipotent cell in the organism, and a few generations of its daughter cells can give rise to all of the cell types in the whole body and the extraembryonic tissues. Pluripotent cells of the blastocyst inner cell mass differentiate to give rise to lineage-committing stem cells and progenitor cells, and finally terminally differentiated cells, gradually losing their differential potential in the process. Pluripotent embryonic stem cells are spun off directly from the inner cell mass of blastocysts and can be cultured *in vitro*. Induced pluripotent stem cells are generated by reprogramming differentiated cells back to the pluripotent state and can also be grown on a dish. Adapted from (Sugawara et al., 2012).

environment (known as "stem cell niche") should be controlled tightly during development, tissue homeostasis or regeneration (Chen et al., 2013; Krause et al., 2013; Morrison and Spradling, 2008; Bendall et al., 2008; Dellatore et al., 2008; Morrison and Kimble, 2006).

Differentiation potential of different stem cells can vary a lot. Development of vertebrates is accompanied by determination of their stem cells, which become less potent and more specialized at every developmental stage. Zygote is the only true totipotent cell, blastomeres are more potent than germ layer stem cells that, in turn can give rise to more cell types than tissue-specific precursors (Mitalipov and Wolf, 2009).

Three types of human stem cells are actively studied today and are believed to have therapeutic potential, namely embryonic stem (ES) cells, adult stem (AS) cells and induced pluripotent stem (iPS) cells.

ES cells

Embryonic stem cells are formed as a normal part of embryonic development. They are isolated from an early embryo (normally from the inner mass of the blastocyst) and can be grown on a dish. ES cells are pluripotent and have the potential to become virtually any cell in human body (Sui et al., 2013; Moon et al., 2011; Kim and Ong, 2012), making them promising for treating many

deceases (Wu et al., 2007; Bernstein, 2012; Lin et al., 2013). At the same time there are a lot of hurdles, both technical and ethical, that researchers developing ES cell therapies face. Delivery of ES cells and their differentiation into desired cell types while avoiding immune response from the patient (Pearl et al., 2012) and tumor formation are only a few of them and not the hardest to overcome. The ethical and legal implications of destroying an embryo in order to isolate ES cells have made some reluctant to support research involving these cells (Patricio, 2011; Nielen et al., 2013).

AS cells

Adult stem cells (also called somatic stem cells) exist naturally in the body and can be found among differentiated cells in different tissues and organs. They are important for growth, healing, and replacing cells that are lost through daily wear and tear. Unlike embryonic stem cells, which are defined by their origin, the origin of adult stem cells in some mature tissues is still under investigation. AS cells have been identified in many organs and tissues, including brain, bone marrow, peripheral blood, blood vessels, skeletal muscle, skin, teeth, heart, gut, liver, ovarian epithelium, and testis (Slack, 2008; Hodgkinson et al., 2009). Most adult stem cells are lineage-restricted, but a number of experiments have reported that certain adult stem cells can trans-differentiate into cell types seen in organs or tissues other than those expected from the cells' predicted lineage (i.e., brain stem cells that differentiate into blood cells or blood-forming cells that differentiate into cardiac muscle cells, and so forth) (Clarke et al., 2000; Krause et al., 2001; Raff, 2003; Passier and Mummery, 2003; Hombach-Klonisch et al., 2008). However, these findings were not properly confirmed, and the ability of adult stem cells to trans-differentiate is still under debate.

The regenerative potential of AS cells drives the current intense interest in adapting them for applications in cell replacement and regeneration therapy. AS cells also don't pose ethical concerns and controversy in comparison with embryonic stem cells, because their extraction and production does not require the destruction of an embryo. Additionally, in instances where adult stem cells are obtained from the intended recipient, the risk of rejection is essentially non-existent. Stem cells from the bone marrow are routinely used as a treatment for blood-related diseases for decades. In 2008 the first full transplant of a human trachea grown from bone marrow stem cells was carried out, and the potential use of these cells to treat other disorders (such as spinal cord injuries or liver cirrhosis) is being intensively investigated (Samuel, 2011; Terai et al., 2006; Subrammaniyan et al., 2011). Therapeutic use of other AS cells is proposed to deal with a vast array of neural, muscle or heart disorders as well as injuries and certain types of cancer (Martinez and Kofidis, 2011; Hernández et al., 2011; Kanno, 2013; Serakinci and Keith, 2006; Mimeault and Batra, 2012; Brunt et al., 2012). Despite recent great efforts to develop AS cell-based therapies, challenges faced by researchers in the field remain numerous. Most types of somatic stem cells are present in low abundance and are difficult to identify, isolate and grow in culture (Paré and Sherley, 2006; Lin et al., 2013; Okano et al., 2013). Isolation itself could cause considerable tissue or organ damage, as in the heart or brain. Self-renewal of AS cells is also not as successful as in embryonic stem cells and as such, they do not proliferate to the same degree. Due to all the same reasons studying the AS cells is challenging just as well. Additionally, functioning of AS cells strongly depends on their microenvironment, their niche (Votteler et al., 2010; Re'em and Cohen,

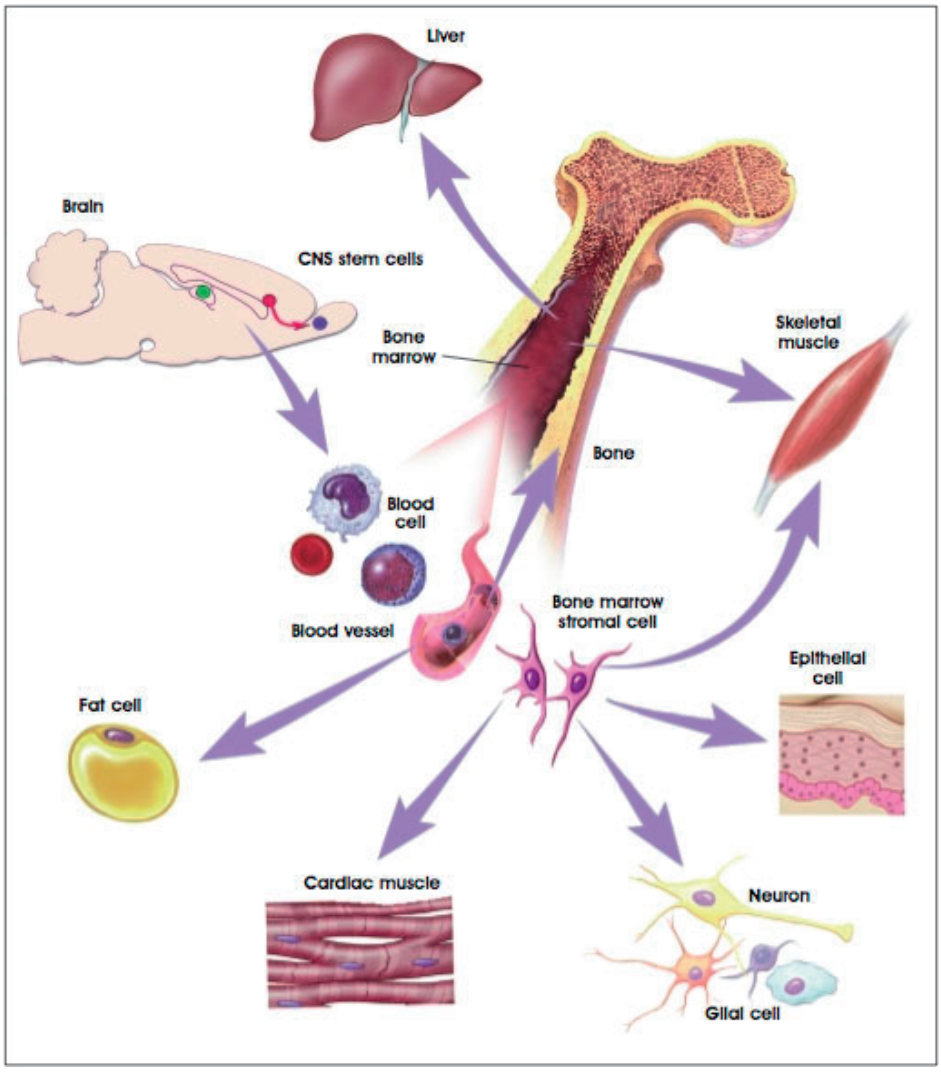


Figure 2. Plasticity of AS cells. Bone marrow stem cells can presumably differentiate into neurons, epithelial and muscle cells, blood stream stem cells give rise to fat cells, and stem cells in the brain could become a source of blood cells (Clarke et al., 2000). Courtesy of Terese Winslow (assisted by Lidia Kibiuk and Caitlin Duckwall).

2011), and the number of models that allow *in vivo* studies of stem cells in the adult organism is still very limited (Gurley et al., 2008).

iPS cells

Induced pluripotent stem cells are adult cells that have been genetically reprogrammed to an embryonic stem cell–like state by being forced to express genes and factors important for

maintaining the defining properties of embryonic stem cells. iPS cells are similar to natural pluripotent stem cells in many aspects, such as expression profiles, chromatin methylation patterns, doubling time, embryoid body formation, teratoma formation, potency and differentiability, but the full extent of their relation to natural stem cells is still being assessed (Takahashi and Yamanaka, 2006; Yu et al., 2007; Kim et al., 2010; Polo et al., 2010). iPS cells have been made from adult stomach and liver cells (Aoi et al., 2008), blood cells (Hanna et al., 2008; Loh et al., 2009), keratinocytes (Aasen et al., 2008), melanocytes (Utikal et al., 2009), pancreatic (Stadtfeld et al., 2008) and urinary tract cells (Moad et al., 2013). iPS cells can become virtually any cell in the body, making them a promising tool for treating many diseases. Using iPS cells is not controversial as compared with ES cells, since no embryos have to be sacrificed in order to collect the cells. In addition, tissues derived from iPSCs will be a nearly identical match to the cell donor and thus probably avoid rejection by the immune system; however, this last assumption was recently challenged (Zhao et al., 2011). iPS cells are already useful tools for drug development and modeling of diseases (Park et al., 2008; Grskovic et al., 2011), and scientists hope to use them in transplantation medicine (Robinton and Daley, 2012; Okano et al., 2013). Recently, the first functional artificial human liver was created using iPS cells (Takebe et al., 2013). Just as in case of other stem cells described, iPS cells may still pose significant risks that could limit their use in humans. It is mostly viruses that are currently used to introduce the reprogramming factors into adult cells, and the expression of oncogenes may be triggered by them (Selvaraj et al., 2010). Alternative methods of reprogramming were proposed, but their efficiency is still very low (Zhou et al., 2009). Besides, even non-viral methods of reprogramming can still lead to DNA damage and rearrangements that, in turn, can trigger tumor formation

Table 1. Comparison of different sources of human stem cells.

	ES cells	**AS cells**	**iPS cells**
Attributes	Pluripotent Easy to identify Easy to extract Easy to grow *in vitro*	Oligopotent Hard to identify Hard to extract Hard to establish cell cultures	Pluripotent Hard to produce Easy to grow *in vitro*
Potential as therapy	Potential source of any cell type in the body	Can be isolated from the patient Blood and bone marrow stem cells are already being used in therapy	Potential source of any cell type Originally can be isolated from the patient
Limitations	Will cause immune response in the patient	Isolation can cause organ damage Low abundance Can cause immune response if not isolated from the patient	Controversial information about possible immune response
Ethical concerns	Human embryos have to be destroyed during isolation Are prohibited in some countries for therapeutic and even research use	Same, as apply to all medical procedures	Same, as apply to all medical procedures

(Knoepfler, 2009; Marión et al., 2009). Another main challenge is controlling the expression of reprogramming factors, since some of them are oncogenes themselves, bringing up the comparison of iPS cells and cancer stem cells (Polo et al., 2010).

Thus, although our knowledge of stem cells is rapidly expanding, and impressive progress in developing regenerative therapies was achieved in the last decade, there are still a lot of gaps to bridge in the field. Many challenges and limits remain, one of the biggest of them being a shortage of ways and approaches to study pluripotent stem cells *in vivo*. Solving this problem would make a big step towards better understanding of stem cell functioning and developing of safer cell transplantation therapies. And one can only make this step with the help of suitable model organism systems.

2. MODEL ORGANISMS USED IN STEM CELL RESEARCH

Stem cells are present in all multicellular organisms. In higher animals of the Animal kingdom, such as vertebrates, pluripotent cells are only present during the embryonic development. During adulthood these animals posses several parallel multipotent systems each responsible for turnover, homeostasis and regeneration of respective tissue (reviewed in (Agata et al., 2006)). Being evolutionary close to human, vertebrate models are extremely useful for development of stem cell-based therapies and clinical trials. The most popular vertebrate models in stem cell research are mice (and much more rarely – rats), amphibians and zebrafish.

Mouse

It is impossible to overestimate the importance of mouse model for stem cell research. Almost all kinds of AS cells were first discovered and identified in mouse, and still most of the AS cell research is done in this model organism (Snippert and Clevers, 2011). ES cells were first extracted from the mouse (Evans and Kaufman, 1981), as well as first iPS cells (Takahashi and Yamanaka, 2006). This is one of the closest models to human evolutionary, and a lot of stem cell-related deceases are studied on mice. Additionally, mouse can form chimeras with transplanted human cells, making it a perfect model for cell therapy and transplantation studies (Steindler, 2007). Tumours in mice also resemble those of humans, so cancer stem cells can also be studied in this organism, as well as the potential of ES, AS or iPS cells to give rise to tumours (Cheng et al., 2010). Importantly, mouse is generally one of the best studied and highly developed model organisms for biomedical research with sophisticated genetic tools and significant genetic information available. Challenges of working with mouse stem cells are almost the same as with human. Pluripotent cells are only available in embryos; they can be extracted and grown *in vitro*, but are hardly accessible for *in vivo* studies. AS cells are only present in small numbers and are hard to identify and study in their microenvironment. Besides, all the general limitations and disadvantages of mouse model apply to stem cell research just the same, namely relatively long reproduction time, amenability to great number of deceases and complicated and expensive lab care. Still, mice remain one of the most powerful models for stem cell research available to date.

Amphibians

Urodele amphibians (axolotls, newts, salamanders) are unmatched in their regeneration abilities among the vertebrates. When injured, these animals can regenerate upper and lower jaw, lens, retina, limb, tail, spinal cord, and even intestine (Brockes and Kumar, 2005). In some cases, the restoration of complex anatomy involves the formation of a blastema, a mass of stem cells that is covered by epithelium and differentiates to replace the missing structures (Chalkley, 1954; Hay and Fischman, 1961). The current data suggest that the blastema may be composed of both de-differentiated somatic cells (Hay and Fischman, 1961; Brockes and Kumar, 2002; Straube and Tanaka, 2006) and reserve stem cells (Morrison et al., 2006), but this has not yet been rigorously elucidated. Urodele amphibians represent an attractive model system for studying stem cell plasticity, migration and possible *in vivo* process of reprogramming differentiated somatic cells into multipotent if not pluripotent state. They are also relatively large animals, making them suitable for transplantation experiments (Stocum, 2000). These animals lay eggs, and that makes their embryos easily accessible for experiments (Beetschen, 1996). Besides, their performance in regeneration (for example, the ability to regenerate spinal cord) already makes urodele amphibians well worth studying. However, we do not yet even know enough about their stem cells to translate findings from this model system to human.

The frog has been a cornerstone of developmental biology for many years. As a research model, this animal has provided researchers with key insights into how a single fertilized egg gives rise to the complex tissues and organ systems of an entire animal (Jones and Smith, 2008). The 2012 Nobel prize in medicine was awarded to John Gurdon for his 1960's experiments involving nuclear transplantation with adult nuclei into frog eggs, these studies were the precursor to current research in stem cells. And today the embryonic development of frogs, especially at early stages, is studied in a great detail (Heasman, 1997; Flachsova et al., 2013). Frogs can regenerate limbs and tails as tadpoles. This regenerative ability is lost and gained during different stages of development and is completely diminished after metamorphosis (Beck et al., 2003; Slack et al., 2004). This suggests that regeneration in frogs may depend upon the presence of undifferentiated cells, which are no longer present once differentiation has set in. Each tissue of the frog tadpole regenerates independently, giving rise to the same tissue (i.e. muscle to muscle, notochord to notochord) (Ryffel et al., 2003; Gargioli and Slack, 2004; Lin et al., 2007), suggesting multipotent stem cells being involved rather than pluripotent. Lately frogs are also used as model for stem cell reprogramming (Gurdon, 2006). Long generation time (1-2 years) and tetraploid genome of the most popular species *Xenopus laevis* are the main, but not the only disadvantages of anuran amphibian (frogs and toads) models.

Zebrafish

Zebrafish has emerged as a powerful model organism for the application of genetics to study vertebrate stem cells during development and regeneration. Embryonic development of these animals is studied in a great detail, a lot is known about gene expression at different stages, and virtually any cell division during early embryogenesis can be tracked (Keller et al., 2008; Vesterlund et al., 2011). Zebrafish can regenerate retina (Bernardos et al., 2007; Nelson et al., 2013), fins (Poss et al., 2003) and heart (Lepilina et al., 2006; Wills et al., 2008). Like with urodele

amphibians, it remains unknown whether zebrafish fin regeneration relies on dedifferentiation of somatic cells or stem/progenitor cell activation, while in the heart multipotent adult stem cells were found that are not only activated by injury, but also constantly contribute to tissue turnover in zebrafish myocardium (Wills et al., 2008). Mutagenesis screens and transcription profiling uncovered a lot of genes associated with stem cell functions in regeneration (Johnson and Weston, 1995; Poss et al., 2002; Lien et al., 2006). By combining forward genetics, pharmacology, transgenic overexpression and recently described targeted mutagenesis (Foley et al., 2009; McCammon et al., 2011) an extremely potent toolbox is now available in zebrafish for a thorough investigation of stem cells and their role in development and regeneration. One of the main technical limitation of this model is a complexity of its genome, which recently went through complete duplication (Postlethwait et al., 2000; Woods et al., 2000).

Drosophila

One of the popular invertebrate models for stem cell research is a fruit fly. *Drosophila melanogaster* retains several populations of stem cells during adulthood as well as transient populations of stem cells during development. Studies of these different populations of stem cells using the genetic tools available to *Drosophila* researchers have played an important role in understanding many conserved stem cell characteristics, such as the role of the niche for maintaining proliferation and undifferentiated state (Pearson et al., 2009; Losick et al., 2011). AS cell populations in *Drosophila* include germline (Spradling et al., 2001; Fuller and Spradling, 2007; Kirilly and Xie, 2007), intestinal (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006), somatic ovary (Margolis and Spradling, 1995; Decotto and Spradling, 2005), hematopoietic (Crozatier and Meister, 2007; Martinez-Agosto et al., 2007), renal and nephric stem cells (Singh et al., 2007). Additionally a population of neural stem cells (also called neuroblasts) is present during development (Doe et al., 1998; Yu et al., 2006). *Drosophila* neuroblasts are the classical model for studies of asymmetric cell division. Different populations of *Drosophila* stem cells are regulated in strikingly different ways. Germ cells stay under the tight control of their microenvironment (so-called stromal cells), executed through physical organization, cell-adhesion and expression of extracellular signals. Only cells capable of responding to the signals from the niche are able to self-renew (Kirilly and Xie, 2007; Fuller and Spradling, 2007). Intestinal stem cells rely much more on the intrinsic signals (Ohlstein and Spradling, 2007; Takashima et al., 2008), while neuroblasts can maintain their stemness without any niche at all (Siegrist and Doe, 2006; Yu et al., 2006). *Drosophila* was also among the first model organisms used to study polarization of cell division, body patterning and body axis formation during development (Nüsslein-Volhard et al., 1987). With a big set of genetic tools developed, fruit fly is a very attractive model for studying multipotent adult stem cells and their regulation, but not pluripotent stem cells.

Models organism to study pluripotent stem cells *in vivo*

All the models mentioned above have a lot of strong sides and advantages, but none of these animals have any pluripotent cells during adulthood (the only exception may be urodele amphibians, but some precise and good-planned experiments still need to be performed to check that). Except for mice, it is hard to speak of pluripotency even during embryogenesis,

since the fate of blastomeres giving rise to the germ line is determined very early in all other models (Extavour and Akam, 2003). Besides, stem cells in higher organisms are relatively difficult to study *in vivo* and not always readily accessible for experimental analysis. Since, in addition, the natural cellular microenvironment is difficult to simulate *in vitro*, the use of alternative model organisms in which fundamental aspects of stem cell biology can be addressed, is highly attractive (Bosch, 2008; Newmark and Sánchez Alvarado, 2002; Tanaka, 2003; Tsai et al., 2002)

Some invertebrates (e.g. Platyhelminthes, Cnidaria, Porifera and Acoela) have astonishing system of adult stem cells that stay pluripotent throughout their life, implicating that a single stem cell population is responsible for tissue homeostasis of the whole organism (Baguñà, 2012; Bode, 1996; Funayama, 2010; Müller, 2006; De Mulder et al., 2009; Gschwentner et al., 2001).

In sea-water living *Hydractinea* (Hydrozoa, Cnidaria), transplantation experiments showed that cnidarian stem cells (so called interstitial cells or I-cells) are able to differentiate into every cell type (Müller et al., 2004; Plickert et al., 2012). Additionally, there are two populations of unipotent cells in Cnidarians that are constantly dividing, giving rise to new ectoderm and endoderm cells. Apart from colonial *Hydractinea* single-living Cnidarians from two different genera are used as model organisms for stem-cell research: fresh-water *Hydra* (Bosch et al., 2010; Galliot, 2012; David, 2012) and recently emerged sea-water *Nematostella* (only one species from this genus, *Nematostella vectensis* is used in the laboratory) (Darling et al., 2005). I-cells are also responsible for regeneration abilities of Cnidarians (first described more than 250 years ago), and today *Hydra* also serves as well-established model for regeneration body axis establishment studies (Trembley, 1744; Sarras, 2012; Galliot, 2013; Bode, 2011).

Sponges (the most commonly used model is *Ephydatia fluviatilis,* Spongellidae, Porifera) have two parallel stem cell systems: pluripotent archaeocytes, that can give rise to any cell in the body, including germ cells, and choanocytes that play role of quiescent stem cells, maintaining the ability to divide and trans-differentiate (Müller, 2006; Funayama, 2010; Funayama, 2013). In 2007 first sponge cell cultures were established, which might add to our knowledge of stem cells in these evolutionary oldest animals (de Caralt et al., 2007).

Stem cell system of planarian flatworms (phylum Platyhelminthes) during adulthood consists of pluripotent somatic cells that can give rise to virtually any cell in the body, including germ line (Sato et al., 2006; Baguñà, 2012). These cells are called neoblasts, they are the only mitotically active cells in the body and constantly proliferate to renew all cell types. Neoblasts are thus required for whole-body homeostasis and are likewise responsible for forming new tissues by growth and regeneration in these animals (Bely and Sikes, 2010; Aboobaker, 2011). Recent studies suggest that certain population of neoblasts can be totipotent (Wagner et al., 2011). We discuss flatworm models and their role in stem cell research in detail in the part 3 of this introduction.

Although unusual among animals in general, a neoblast stem cell system is also known from a group of small, soft-bodied marine worms known as acoels (Bely and Sikes, 2010). Stem cell system of acoel species *Isodiametra pulchra* was recently described (De Mulder et al., 2009). Generally, acoels were considered members of the flatworms (Platyhelminthes) based on a combination of weak morphological characters. However recently presented molecular and

developmental data makes the Acoela a completely separate group at the very base of Bilateria (Egger et al., 2009), making the similarity between stem cell systems of flatworms and animals like *I. pulchra* very striking. Based on these data tissue turnover and regeneration mechanisms present in flatworms might appear evolutionary old and well-conserved. More recently, the basal phylogenetic position of Acoels have been challenged (Philippe et al., 2011) and according to this work Acoels might be much closer to chordates than previously thought, which makes the investigation of stem cell system in Acoels even more interestring.

Presence of an adult pluripotent stem cell system within these four taxa (Platyhelminthes, Cnidaria, Porifera and Acoela) could be indirectly linked to the ability of asexual reproduction (Agata et al., 2006). In planarians, for example, asexual reproduction typically takes place by fission and neoblasts are responsible for reconstructing missing parts of the body in both "daughter" animals (Saló, 2006). In sponges small particles called gemmules are responsible for asexual reproduction, and these primordial gemmules are formed by the divisions of pluripotent archaeocytes (Müller, 2006; Funayama, 2013).

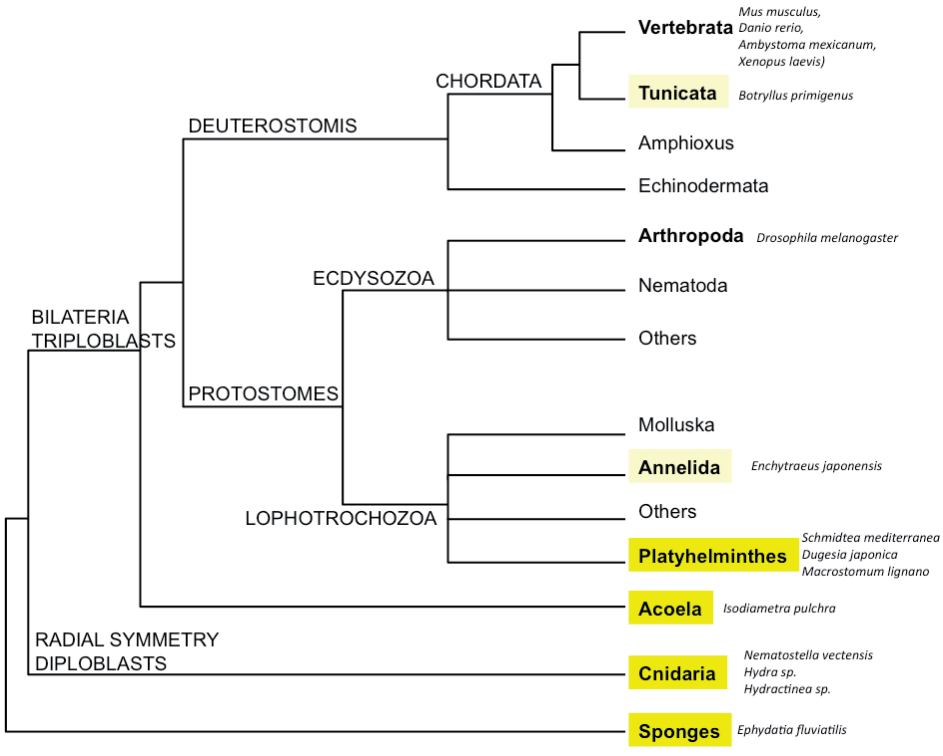


Figure 3. Simplified phylogenetic tree of the Animal kingdom. Species used as model organisms for stem cell research are shown and taxa they belong to are highlighted in bold font. Pluripotent somatic stem cells, capable of differentiating into all cell types, including germ line, are shown in yellow background. Annelida and Tunicata are shown in bright-yellow, since the existence of such stem cell populations in the species from these taxa is not yet clear. Modified from (Adoutte et al., 2000).

In all the animals possessing adult pluripotent stem cell systems these cells can give rise to the germ cells (Agata et al., 2006). Germ line deficient fragments of flatworms can give rise to the whole animal with functional germ cells (Morgan, 1902; Sánchez Alvarado, 2006; Handberg-Thorsager et al., 2007; Rink, 2013). In *Hydra* germ cells differentiate from I-cells during a process of sexualisation (Mochizuki et al., 2000; Mochizuki et al., 2001).

Similar situation was described in some ascidian (urochordate) model species, though generally ability to produce germ cells from somatic stem cells is not common for deuterostomes. Interestingly, these very species, colonial ascidians, are capable of asexual reproduction by budding (Kawamura et al., 2008; Nakauchi, 1982) and demonstrate impressive regeneration abilities (Tiozzo et al., 2008). In *Botryllus primigenus* somatic stem cells called hemoblasts are found in the blood stream (Sugino et al., 2007) and can produce not only somatic cells (Satoh, 1994), but also give rise to gonads and germ line (Sunanaga et al., 2006; Kawamura and Sunanaga, 2010). Alternative hypothesis, however, postulates that there are two independent hemoblast lineages, separated early in embryonic development, one giving rise to somatic cells, and the other to germ line (Laird et al., 2005; Brown et al., 2009). More experiments need to be performed to make the origins and properties of hemoblasts clear.

In some annelids capable of asexual reproduction and extensive regeneration, like *Enchytraeus japonensis* (Oligochaeta, Annelida) (Sugio et al., 2012) somatic and germ stem cell lineages are separated early during embryogenesis and somatic stem cells cannot give rise to germ line (Sugio et al., 2008; Tadokoro et al., 2006; Yoshida-Noro and Tochinai, 2010), suggesting that pluripotent stem cells in the adults that are capable of differentiating into any cell type, are only present in more basal organisms.

Evolution of stem cell regulation

Each model system for stem cell research has its own properties, sometimes unique, and its own advantages and disadvantages. The field can benefit from the integration of the molecular and cellular knowledge gathered from all these different organisms. By comparing different stem cell populations within different model organisms, we can get a true understanding of how stem cell regulation has evolved and how does it work. But how conservative are the molecular mechanisms responsible for such regulation in different taxa?

Evolutionary conservation of the same signalling pathways in distinct stem cell systems has only been studied for a little more than a decade (Benfey, 1999; Zhan et al., 2005). Large scale species comparative studies at the genome level have revealed that early branching metazoans such as sponges, cnidarians and acoels possess many if not most of stem cell specific genes characteristic for vertebrates including human (Benfey, 1999; Bosch, 2008; Hemmrich and Bosch, 2008). Intriguingly though, out of the three nuclear factors (*Oct-4*, *Sox-2* and *Nanog*) that govern pluripotency *in vivo* and *in vitro* in vertebrates (Pan and Thomson, 2007; Yu et al., 2007; Silva and Smith, 2008; Mullen et al., 2011), only *Sox-2* could be found in the genomes of lower metazoans (Hemmrich and Bosch, 2008). Being present in different taxa stem cell specific genes may still have different functions, and signaling pathways do not necessarily act the same way in different stem cell systems. Some stem cell related signaling cascades, including *Wnt* signaling and *Hippo* pathway appear to be mostly conserved on the functional level (Sato et

al., 2004; Almuedo-Castillo et al., 2012; Holland et al., 2013; Lian et al., 2010; Tamm et al., 2011; Demircan and Berezikov, 2013; Barry and Camargo, 2013). At the same time homologues of the key players of *TGF-β* and *FGFR* pathways, involved in maintaining pluripotency in human ES cells (James et al., 2005; Vallier et al., 2005) and *BMP* pathway, responsible for differentiation in both ES and AS cells (Li and Chen, 2013) can be found in basal model organisms, but their roles are not always same and to large extend remain unclear. In mouse ES cells *BMP* signaling maintains the self-renewal state (Li et al., 2012; Ying et al., 2003) in contrast to human. *Notch* pathway is important for intestinal stem cell self-renewal (Vanuytsel et al., 2013), whereas it does not play a significant role in ES cells pluripotency (Noggle et al., 2006). Thus it should be kept in mind that on the functional level significant differences in the signaling pathways can be found between very evolutionary close species and even between different stem cell populations of the same organism.

This common lack of functional conservation together with our incomplete understanding of the signaling pathways that regulate different stem cell populations make it impossible to identify universal molecular markers of stem cells. For example, expression of *Piwi* and *Vasa* proteins that are highly conserved in all animals and even plants (in case of *Piwi*), is characteristic for the germ stem cells and their niches in most of the model organisms, and their knockout or knockdown results in partial or complete sterility (Cox et al., 2000; Cox et al., 1998; Houwing et al., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004; Juliano et al., 2011; Hay et al., 1988; Lasko and Ashburner, 1988; Fujiwara et al., 1994; Komiya et al., 1994; Yoon et al., 1997; McDougall et al., 2011). At the same time in flatworms, acoels, sponges and cnidarians expression of both proteins is extended to pluripotent somatic stem cells and is not specific for germ line development anymore (Denker et al., 2008; Reddien et al., 2005; Seipel et al., 2004; De Mulder et al., 2009; Funayama et al., 2010; Mochizuki et al., 2001; Pfister et al., 2008; Shibata et al., 1999). On the other hand, cell division markers that are expressed in a great number of cell populations in higher organisms can be routinely used to identify pluripotent stem cells and germ cells in flatworms. In planarians *PCNA, CyclinB, H2b* and *MCM2* were used to localize stem cells and follow stem cell dynamics during different biological processes (Solana et al., 2012; Orii et al., 2005; Salvetti et al., 2000).

In conclusion, although data obtained from different model systems in the last two decades remarkably extended our understanding of stem cell functioning, our knowledge of the mechanisms regulating both proliferation and differentiation of stem cells still remains limited. Deep comparative analysis of stem cell populations and their functioning in diverse organisms promises new insights into processes of development and maintenance of different tissues conducted by stem cells. Additionally, such analysis would reveal how diverse stem cell systems may have evolved. Comparative studies performed on different model organisms and exploration of new suitable models, especially the ones possessing accessible pluripotent stem cell populations, would contribute greatly to our understanding of fundamental mechanisms regulating stem cells and would potentially lead to development of safe and efficient stem cell based therapies.

One of the most attractive and well-studied groups of animals with pluripotent stem cells available during adulthood is flatworms, and it is discussed in more detail below.

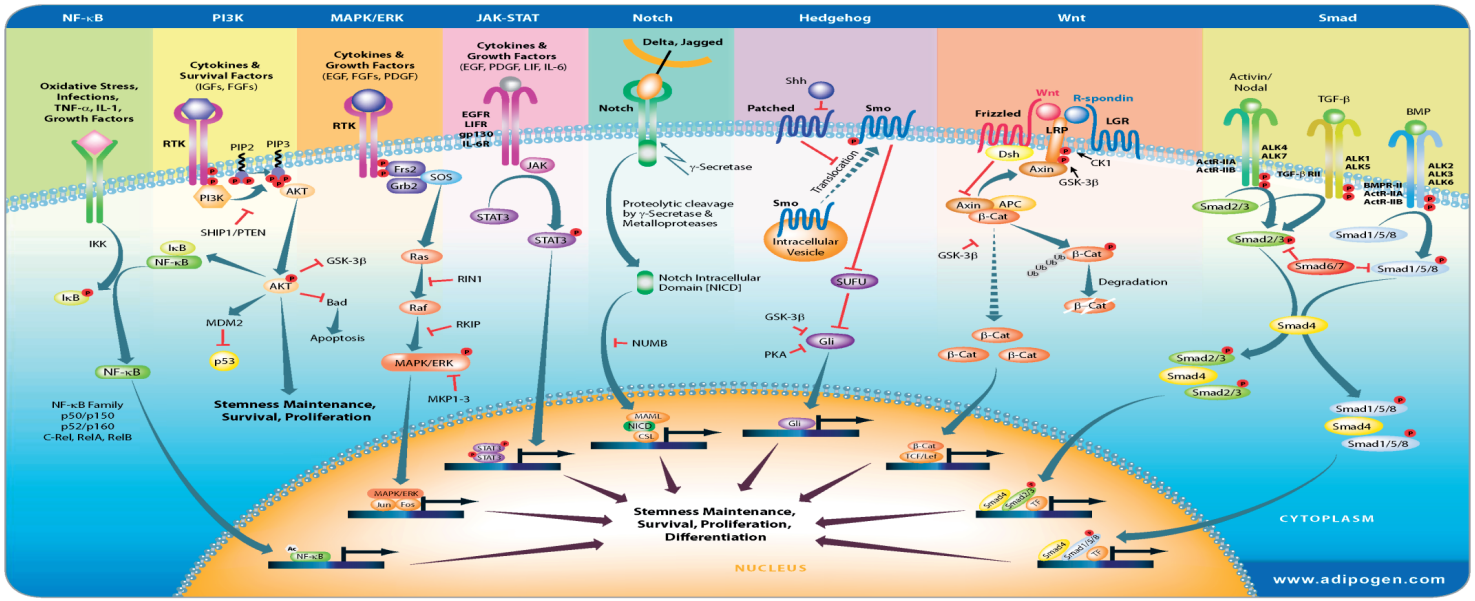


Figure 4. Conserved signalling pathways involved in stem cell regulation. All the pathways depicted and most of their members appear to be highly conserved between different animal species. All of them regulate proliferation, stemness maintenance or differentiation of the cells. However, in different stem cell populations the importance of the same pathways might be different, and the same pathway members can play opposite roles (examples in the text). From Cancer & Stem Cell Signalling Pathways, edition April 2013, at www.adipogen.com

3. FLATWORM MODELS

Flatworms (Platyhelminthes) are one of the simplest animals that are bilaterally symmetrical, lack the body cavity and specialized respiratory and circulatory system (Rieger et al., 1991). In addition, flatworms possess a primitive cellular mesoderm and an anterior concentration of neural tissue. The gut is separated from the epidermis by mesenchymal cells (called parenchyma) (Brusca and Brusca, 2003). Phylum Platyhelminthes belongs to Lophotrochozoa, the most basal bilaterian animals (Edgecombe et al., 2011).

Since the end of the 19th century, when regeneration abilities of the species *Planaria maculata* and *Planaria lugubris* had been discovered and described by Thomas Hunt Morgan (Morgan, 1898; Morgan, 1902), planarian flatworms have been used by a great number of groups, becoming in the late 20th century a traditional model for stem cell and regeneration research. Today several flatworms species are used in laboratories, most of them belonging to the order Tricladida (so called planarian flatworms), like *Schmidtea mediterranea*, *Dugesia japonica, Schmidtea polychroa, Girardia tigrina* and others (Newmark and Sánchez Alvarado, 2002; Sánchez Alvarado, 2004; Agata et al., 2006; Saló, 2006; Gentile et al., 2011). The only non-planarian free-living flatworm model that is lately being routinely used by a number of laboratories is *Macrostomum lignano* (order Macrostomida) (Ladurner et al., 2005).

Planarian flatworms are small and are easy to culture in the laboratory conditions. They grow fast and can reproduce asexually, which provides a great opportunity to create and expand population of genetically identical animals in a short time (Sánchez Alvarado, 2004; Gentile et al., 2011) But it is their impressive regeneration potential, facilitated by adult pluripotent stem cell system, what makes them truly special and incredibly attractive as a model for stem cell research (Reddien and Sánchez Alvarado, 2004; Agata et al., 2006; Rink, 2013).

Neoblasts of planarian flatworms are accessible to manipulations within their micro-environment and thus can be studied *in vivo* (Sánchez Alvarado, 2007). Stem cell dynamics during post-embryonic development and in response to amputation (Egger et al., 2007) or starvation (González-Estévez and Saló, 2010; Pellettieri and Sánchez Alvarado, 2007) can be addressed in this model.

Planarian neoblasts were extensively studied by a number of laboratories, especially in the last twenty years, and knowledge collected during these years is summarized in several recent reviews (Sánchez Alvarado, 2007; Aboobaker, 2011; Baguñà, 2012; Rink, 2013). Neoblasts are small undifferentiated adult cells that have capacity to give rise to various cell types while retaining self-renewal potential thereby fulfilling the definition of stem cells (Rieger et al., 1999; Sánchez Alvarado and Kang, 2005; Baguñà, 2012; Rink, 2013). Neoblasts are the only mitotically active cells (Morita and Best, 1984; Baguñà, 2012) in planarian flatworms. Accumulated data from light and electron microscopy and cell lineage tracing experiments support capacity of neoblasts to turn into all cellular types, including germ line (Sato et al., 2006). Recent single cell transplantation experiments showed that some planarian neoblasts are likely totipotent (Wagner et al., 2011). Cell sorting combined with gene expression analysis indicate that neoblast population is likely to be heterogeneous (Hayashi et al., 2006; Higuchi et al., 2007; Shibata et al., 2012). This hypothesis is further supported by X-ray treatment experiments: two sub-groups of neoblasts (X1 and X2 subgroups) can be separated based on their response to X-ray

irradiation (Salvetti et al., 2009). Morphologically there are some minor differences between different neoblasts as well, but generally they appear as small (6-10µm) cells with large nucleus and prominent nucleolus and only a thin rim of cytoplasm with free ribosomes, chromatoid bodies and few mitochondria (Pedersen, 1959; Hori, 1997; Rieger et al., 1999; Baguñà, 2012).

During the last 15 years fast progress and expansion of stem cell field together with development of new techniques and experimental methods have sparked planarian research. The genome of *S. mediterranea* is available (Robb et al., 2008) and a genome project of *D. japonica* has been initiated. The current planarian toolkit includes robust in situ hybridization (ISH) (Umesono et al., 1997; Pearson et al., 2009), BrdU–labeling (Newmark and Sánchez Alvarado, 2000) and cell sorting (Hayashi et al., 2006; Hayashi and Agata, 2012) that are routinely used in many laboratories. RNA interference (RNAi) experiments (Sánchez Alvarado and Newmark, 1999; Newmark et al., 2003; Reddien et al., 2005) made it possible to identify signal transduction pathways and specific genes that regulate different aspects of regeneration (Gurley et al., 2008; Petersen and Reddien, 2008; Adell et al., 2009; Rink et al., 2009; Felix and Aboobaker, 2010; Oviedo et al., 2008) and stem cell functioning (Oviedo et al., 2008; Guo et al., 2006; Fernandéz-Taboada et al., 2010; Scimone et al., 2010). Planarian neoblasts can be specifically eliminated (completely or partially) by X-ray irradiation (Wolff, 1948), and this technique is widely used to prove stem cell specific expression of certain genes as well as for identification of novel genes important for neoblast functioning (Orii et al., 2005; Reddien et al., 2005; Rossi et al., 2006; Eisenhoffer et al., 2008) Lethal doses of γ-irradiation can be used alternatively for the same purposes (Wagner et al., 2011; Solana et al., 2012).

Combination of all the methods mentioned above with next generation sequencing techniques made planarians an outstanding tool for gene discovery (Friedländer et al., 2009; Solana et al., 2012; Onal et al., 2012). Several independent transcriptome assemblies have been completed lately (Blythe et al., 2010; Adamidi et al., 2011; Sandmann et al., 2011; Nishimura et

Figure 5. Examples of planarian species. From left to right: *Polycelis sp., Planaria torva, Dendrocoelum lacteum, Schmidtea polychroa, Dugesia gonocephala, Schmidtea mediterranea*. From (Rink, 2013).

al., 2012), and different approaches, combining RNA-seq with irradiation or RNAi treatments and/or cell sorting have been used to identify genes involved in proliferation and regulation of planarian stem cells (Friedländer et al., 2009; Solana et al., 2012; Onal et al., 2012; Blythe et al., 2010; Rossi et al., 2007; Eisenhoffer et al., 2008; Wagner et al., 2012). Several versions of potential "neoblast-specific" transcriptomes were created and among many genes now known to be expressed in neoblasts, one can pinpoint a few groups such as components of the cell division machinery, proteins involved in posttranscriptional regulation of gene expression, chromatin modifiers (Solana et al., 2012; Onal et al., 2012; Rossi et al., 2007; Eisenhoffer et al., 2008; Wagner et al., 2012; Shibata et al., 2012; Labbé et al., 2012; Rouhana et al., 2010).

Planarian flatworm model system contributed a lot to our understanding of stem cell functioning, yet there are limits at the moment to what we could learn from these animals. There are technical challenges in the field that seem hard to overcome at the moment. Transgenics are still not available for planarian flatworms, limiting their use as model organisms greatly, and struggles with sexual reproduction, characteristic for these animals, together with their ability to reproduce asexually make it very hard to use forward genetics approaches in planarian flatworms. Despite all the recent advances in understanding neoblast biology, a number of unanswered questions remain in the field. Experiments of Wagner and colleagues demonstrated that at least 5% of planarian neoblasts are indeed pluripotent, but also that some are clearly not (Wagner et al., 2011). This unknown degree of population heterogeneity amongst neoblasts (supported by other experiments mentioned above) remains one of the biggest problems in planarian stem cell research, both conceptually and experimentally (Baguñà, 2012; Rink, 2013). We cannot yet define these stem cell subpopulations and still know little about the mechanisms that regulate dynamic steady state between different populations of neoblasts and differentiated cells. Almost nothing is known about embryonic origins of adult somatic stem cells, regulation of neoblast divisions at the single cell level (Baguñà, 2012; Rink, 2013) and stage of planarian embryonic or post-embryonic development, when different subpopulations of stem cells become separated. More information, detailed transcription profiling of stem cell subpopulations and, most importantly, new tools that are not yet available in planarian flatworms, such as transgenics, would be required to resolve these questions. Another intriguing question concerning neoblasts is evolution of pluripotency. According to some recently acquired data, there is a high degree overlap between sets of genes expressed in neoblasts and vertebrate stem cells (Onal et al., 2012; Labbé et al., 2012). Comparison of pluripotency on the molecular level across shorter evolutionary distances would contribute to our understanding of stem cell systems evolution and potentially would help us to understand the regulation of planarian neoblasts better.

Growing interest in new flatworm model systems is reflected in three recent studies. Three different groups used new planarian species, *Phagokata kawakatsui*, *Procotyla fluviatilis* and *Dendrocoelum lacteum* in their studies of the Wnt signaling and its role in regeneration (Umesono et al., 2013; Sikes and Newmark, 2013; Liu et al., 2013).

One of the non-planarian organisms that is relatively close to triclads evolutionary and morphologically, yet has some significant differences, is flatworm *Macrostomum lignano* (order Macrostomida) that recently emerged as a new model for stem cell research (Ladurner et al., 2005).

4. *MACROSTOMUM LIGNANO*

M. lignano is sea water free-living flatworm, about 1,5 mm long, that consists of roughly 25000 cells. *M. lignano* represents a basal member of free-living worms within the Rhabditophora, the largest taxon within the phylum Platyhelminthes. It emerged as a model organism less than a decade ago (Ladurner et al., 2005) but today *M. lignano* is used in a number of laboratories (that keeps on growing every year) and recently has proven itself as a useful invertebrate model for stem cell, regeneration, aging and sex allocation research (Egger et al., 2006; Pfister et al., 2008; De Mulder et al., 2009; Mouton et al., 2009; Verdoodt et al., 2012; Simanov et al., 2012; Janicke et al., 2013; Demircan and Berezikov, 2013).

M. lignano populations are easy to grow and maintain in laboratory conditions, and several cultures are kept in the laboratory for many years. Worms are cultured in Petri dishes in f/2 medium (nutrient-enriched artificial sea water), and are fed *ad libitum* with diatom *Nitzchia curvilineata*. Distinct morphology with clearly formed and visible organs (gonads, reproductive organs, brain, various glands) together with transparency of the worm makes it possible to

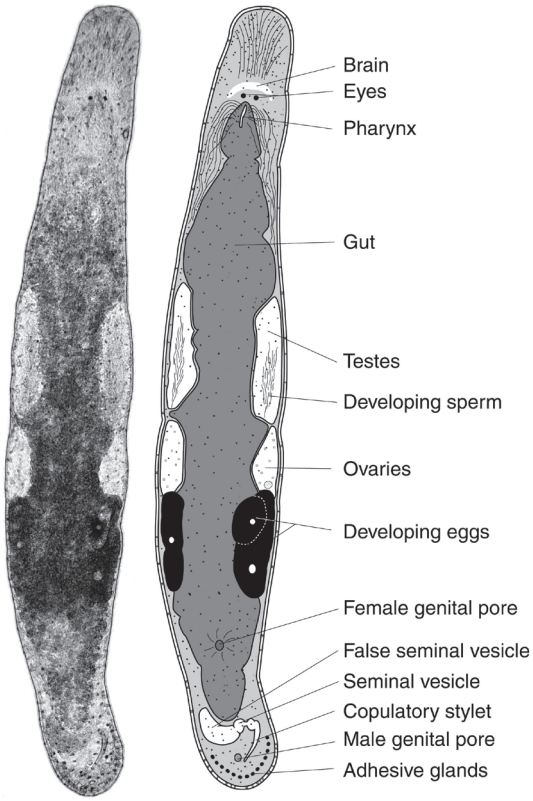


Figure 6. DIC-image and the drawing of adult *Macrostomum lignano*. Pharynx and gut, eyes and brain, organs of reproductive system and adhesive glands are easy to identify in the transparent worm. Courtesy of Lukas Schärer.

monitor morphological responses to different conditions or treatments in a living specimen under the microscope (Nimeth et al., 2002; Nimeth et al., 2004; Pfister et al., 2007; De Mulder et al., 2009). Besides, different fluorescent and non-fluorescent labelings can be performed without additional bleaching steps (Ladurner et al., 2005; Pfister et al., 2007; Pfister et al., 2008). *M. lignano* is an obligatory non-self fertilizing hermaphrodite and has exclusively sexual reproduction. Well-fed adult animals generate a lot of embryos all year through (one animal lays one egg a day on average), making it amenable for genetic manipulations (Demircan et al., in preparation). Unlimited access to single eggs during the whole year also allows the analysis of the embryonic development (Morris et al., 2004; Morris et al., 2006).

Dividing cells can be visualized in S-phase by BrdU labeling (that is executed by soaking) and in M-phase by pH3-labeling, which are crucial to follow stem cell dynamics during different treatments or biological processes (Nimeth et al., 2004; Pfister et al., 2007; Demircan and Berezikov, 2013). Current toolkit available for work with *M. lignano* further includes robust ISH technique (Pfister et al., 2007; Pfister et al., 2008), set of tissue-specific monoclonal antibodies (Ladurner et al., 2005) and simple RNAi procedure by soaking (Pfister et al., 2008; Demircan and Berezikov, 2013). The draft genome and transcriptome of *M. lignano* are now sequenced and annotated (see http://www.macgenome.org), and, most importantly, transgenics, including site-specific mutagenesis, are available (Demircan et al, in preparation). Just like in planarian flatworms, γ-irradiation can be used to specifically eliminate neoblasts, however, much higher doses are used since stem cells of *M. lignano* demonstrate remarkably high resistance to irradiation (De Mulder et al., 2010).

Table 2. Experimental properties of *M. lignano* in comparison with *S. mediterranea*.

	Macrostomum lignano	***Schmidtea mediterranea***
Size	1 mm	> 5 mm
Total cell number	+/– 25'000	Millions
Neoblasts	+/– 1600 (6–7%)	25–35%
Transparency	Highly transparent	Opaque
Culturing media	f/2 (sea water based)	Pond or spring water (fresh water)
Feeding	Diatom algae (*Nitzschia curvilineata*)	Beef liver
Reproduction	Only sexual	Mainly asexual, seasonally sexual
Embryogenesis	5 days	Several weeks
Generation time	18 days	Months
Nervous-, muscle systems and gonads	Simple	Complex
Stem Cell system	Pluripotent	Pluripotent
BrdU/H3 staining	By soaking	By feeding or injection
RNA interference	By soaking	By feeding or injection
Genome	Publicly available	Publicly available
Accessibility to eggs	Single eggs whole year through	Eggs in cocoons, available seasonally
Transgenics	Established, by injection into the egg	Not available

Neoblasts represent around 6-7 % of all the cells in adult worm. Just like in planarian flatworms they are involved in tissue turnover and also facilitate regeneration capacity of the animal. Being capable to regenerate almost any part of the body, *M. lignano* is still unable to regenerate the head. Posterior-facing blastemas give origin to fully functioning tails with all organs and structures, whereas anterior-facing wounds develop blastema layer but no actual regeneration happens (Egger et al., 2006). This difference in regeneration capacity of *M. lignano* and planarian model species raises very intriguing questions. *Macrostomum* neoblasts are capable of giving rise to head-specific structures like eyes, pharynx or brain (Egger et al., 2006; Simanov et al., 2012), so there must be certain regulation differences responsible for the limitations of *M. lignano* regeneration capacity. Identifying these differences would contribute a lot to our understanding of stem cells and specifically their role in regeneration.

All the neoblasts in *M. lignano* are located mesodermally, mainly on the sides of the body and never anterior from the eye level. A small number of neoblasts also reside around the middle line at the pharynx or tail plate. Additionally, proliferating cells can be found in gonads, and somatic stem cells can give origin to germ line (Rieger et al., 1999; Ladurner et al., 2000; Bode et al., 2006). Neoblasts look alike stem cells in planarian flatworms. They are characterized by small size (5-10 µm) and round shape with high nuclear-cytoplasmic ratio (Rieger et al., 1999; Bode et al., 2006).

In conclusion, *M. lignano* is a promising model system for *in vivo* stem cell biology. It can contribute to our understanding of fundamental mechanisms regulating stem cells and evolution of pluripotency. It can also complement the research done on planarian worms and add to our knowledge of neoblasts. We are convinced that *M. lignano* can also be used successfully in other research fields, and number of questions that can be addressed in this model system is by far not limited by the properties of stem cells. But in order to conduct all kinds of research one first needs a reliable toolkit and knowledge of the model system including its genomic resources. That is what this thesis aims to provide.

5. THESIS AIM AND OUTLINE

The stem cell research field today is short of suitable model systems, especially the ones allowing *in vivo* studies of pluripotent cells. Our knowledge of stem cells expanded greatly in the last decennia, and today the humanity stands just one step away from developing safe and reliable stem cell based therapies to fight a great number of deceases, degenerative disorders and probably the aging itself. However, we still cannot make this step, because knowing a lot about different stem cell populations we still lack clear understanding of fundamental mechanisms that regulate pluripotency and controlled differentiation. Exploring new model systems would add more pieces to this puzzle, bringing it closer to the nice and clear-cut picture we want to see.

This thesis summarizes a part of work that was done by our group in the last 4 years in order to develop *Macrostomum lignano* as a reliable model system for stem cell research.

Chapter 1 has given a general brief introduction into the state of art in today's stem cell research. We focused on the model organisms that are currently used in the field and described stem cell systems of each of them. We also discussed a great potential of flatworm stem cell

models and introduced *M. lignano* as a new model species that would become complimentary to existent planarian models.

4 years ago genomic resources of *M. lignano* were limited to EST database and a few cloned genes. In the **Chapter 2** we present a draft genome of the species, which we sequenced, assembled and annotated.

Chapter 3 adds another level to our knowledge of the genomic properties of *M. lignano*. We present first *de novo* assembled transcriptome of *M. lignano*. To that we added RNA-seq data collected from irradiated (neoblast-ablated) worms and animals at different developmental stages. We compared these sets of RNA-seq data and identified a number of stem cell specific genes in *M. lignano*.

M. lignano lives in the water and is therefore amenable to treatment by different compounds by soaking. In **Chapter 4** we show how by using DiBAC dyes we can visualize membrane potentials and stable patterns they form in the whole animal *in vivo*. Disruption of these patterns by ion channel targeting drugs leads to changes in regeneration polarity and can possibly induce the head regeneration.

Chapter 5 deals with high resistance of *M. lignano* to the changes in water salinity and response of neoblasts and germ line cells to extremely low salinity conditions. Gene expression dynamics after changing the salinity of the media is also addressed.

In **Chapter 6** we return to the topic of genomic recourses of *M. lignano*. We describe microRNAs that we found in its genome and unusual structure of their precursors that appear characteristic for this animal.

Chapter 7 gives a summary of this thesis. We discuss the results of all the studies mentioned above, state of art of ongoing projects and possible directions for future studies.

REFERENCES

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotechnol *26*, 1276-284.

Aboobaker, A.A. (2011). Planarian stem cells: a simple paradigm for regeneration. Trends Cell Biol *21*, 304-311.

Adamidi, C., Wang, Y., Gruen, D., Mastrobuoni, G., You, X., Tolle, D., Dodt, M., Mackowiak, S.D., Gogol-Doering, A., et al. (2011). De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics. Genome Res *21*, 1193-1200.

Adell, T., Salò, E., Boutros, M., and Bartscherer, K. (2009). Smed-Evi/Wntless is required for beta-catenin-dependent and -independent processes during planarian regeneration. Development *136*, 905-910.

Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'homme, B., and de Rosa, R. (2000). The new animal phylogeny: reliability and implications. Proc Natl Acad Sci U S A *97*, 4453-56.

Agata, K., Nakajima, E., Funayama, N., Shibata, N., Saito, Y., and Umesono, Y. (2006). Two different evolutionary origins of stem cell systems and their molecular basis. Semin Cell Dev Biol *17*, 503-09.

Agata, K., Nakajima, E., Funayama, N., Shibata, N., Saito, Y., and Umesono, Y. (2006). Two different evolutionary origins of stem cell systems and their molecular basis. Semin Cell Dev Biol *17*, 503-09.

Almuedo-Castillo, M., Sureda-Gómez, M., and Adell, T. (2012). Wnt signaling in planarians: new answers to old questions. Int J Dev Biol *56*, 53-65.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science *321*, 699-702.

Baguñà, J. (2012). The planarian neoblast: the rambling history of its origin and some current black boxes. Int J Dev Biol *56*, 19-37.

Barry, E.R., and Camargo, F.D. (2013). The Hippo superhighway: signaling crossroads converging on the Hippo/Yap pathway in stem cells and development. Curr Opin Cell Biol *25*, 247-253.

Beck, C.W., Christen, B., and Slack, J.M. (2003). Molecular pathways needed for regeneration of spinal cord and muscle in a vertebrate. Dev Cell *5*, 429-439.

Beetschen, J.C. (1996). How did urodele embryos come into prominence as a model system? Int J Dev Biol *40*, 629-636.

Beltrami, A.P., Cesselli, D., and Beltrami, C.A. (2011). At the stem of youth and health. Pharmacol Ther *129*, 3-20.

Bely, A.E., and Sikes, J.M. (2010). Acoel and platyhelminth models for stem-cell research. J Biol *9*, 14.

Bendall, S.C., Stewart, M.H., and Bhatia, M. (2008). Human embryonic stem cells: lessons from stem cell niches in vivo. Regen Med *3*, 365-376.

Benfey, P.N. (1999). Stem cells: A tale of two kingdoms. Curr Biol *9*, R171-72.

Bernardos, R.L., Barthel, L.K., Meyers, J.R., and Raymond, P.A. (2007). Late-stage neuronal progenitors in the retina are radial Müller glia that function as retinal stem cells. J Neurosci *27*, 7028-040.

Bernstein, H.S. (2012). Cardiac repair and restoration using human embryonic stem cells. Regen Med *7*, 697-712.

Blythe, M.J., Kao, D., Malla, S., Rowsell, J., Wilson, R., Evans, D., Jowett, J., Hall, A., Lemay, V., et al. (2010). A dual platform approach to transcript discovery for the planarian Schmidtea mediterranea to establish RNAseq for stem cell and regeneration biology. PLoS One *5*, e15617.

Bode, A., Salvenmoser, W., Nimeth, K., Mahlknecht, M., Adamski, Z., Rieger, R.M., Peter, R., and Ladurner, P. (2006). Immunogold-labeled S-phase neoblasts, total neoblast number, their distribution, and evidence for arrested neoblasts in Macrostomum lignano (Platyhelminthes, Rhabditophora). Cell Tissue Res *325*, 577-587.

Bode, H. (2011). Axis formation in hydra. Annu Rev Genet *45*, 105-117.

Bode, H.R. (1996). The interstitial cell lineage of hydra: a stem cell system that arose early in evolution. J Cell Sci *109 (Pt 6)*, 1155-164.

Bosch, T.C. (2008). Stem cells (Springer).

Bosch, T.C., Anton-Erxleben, F., Hemmrich, G., and Khalturin, K. (2010). The Hydra polyp: nothing but an active stem cell community. Dev Growth Differ *52*, 15-25.

Brockes, J.P., and Kumar, A. (2002). Plasticity and reprogramming of differentiated cells in amphibian regeneration. Nat Rev Mol Cell Biol *3*, 566-574.

Brockes, J.P., and Kumar, A. (2005). Appendage regeneration in adult vertebrates and implications for regenerative medicine. Science *310*, 1919-923.

Brown, F.D., Tiozzo, S., Roux, M.M., Ishizuka, K., Swalla, B.J., and De Tomaso, A.W. (2009). Early lineage specification of long-lived germline precursors in the colonial ascidian Botryllus schlosseri. Development *136*, 3485-494.

Brunt, K.R., Weisel, R.D., and Li, R.K. (2012). Stem cells and regenerative medicine - future perspectives. Can J Physiol Pharmacol *90*, 327-335.

Brusca, R.C., and Brusca, G.J. (2003). Invertebrates (Sunderland, Mass.: Sinauer Associates).

de Caralt, S., Uriz, M.J., and Wijffels, R.H. (2007). Cell culture from sponges: pluripotency and immortality. Trends Biotechnol *25*, 467-471.

Chalkley, D.T. (1954). A quantitative histological analysis of forelimb regeneration in Triturus viridescens. Journal of Morphology *94*, 21-70.

Chen, S., Lewallen, M., and Xie, T. (2013). Adhesion in the stem cell niche: biological roles and regulation. Development *140*, 255-265.

Cheng, L., Ramesh, A.V., Flesken-Nikitin, A., Choi, J., and Nikitin, A.Y. (2010). Mouse models for cancer stem cell research. Toxicol Pathol *38*, 62-71.

Clarke, D.L., Johansson, C.B., Wilbertz, J., Veress, B., Nilsson, E., Karlström, H., Lendahl, U., and Frisén, J. (2000). Generalized potential of adult neural stem cells. Science *288*, 1660-63.

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

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev *12*, 3715-727.

Crozatier, M., and Meister, M. (2007). Drosophila haematopoiesis. Cell Microbiol *9*, 1117-126.

Darling, J.A., Reitzel, A.R., Burton, P.M., Mazza, M.E., Ryan, J.F., Sullivan, J.C., and Finnerty, J.R. (2005). Rising starlet: the starlet sea anemone, Nematostella vectensis. Bioessays *27*, 211-221.

David, C.N. (2012). Interstitial stem cells in Hydra: multipotency and decision-making. Int J Dev Biol *56*, 489-497.

Decotto, E., and Spradling, A.C. (2005). The Drosophila ovarian and testis stem cell niches: similar somatic stem cells and signals. Dev Cell *9*, 501-510.

DeGregori, J. (2011). Evolved tumor suppression: why are we so good at not getting cancer? Cancer Res *71*, 3739-744.

Dellatore, S.M., Garcia, A.S., and Miller, W.M. (2008). Mimicking stem cell niches to increase stem cell expansion. Curr Opin Biotechnol *19*, 534-540.

Demircan, T., and Berezikov, E. (2013). The Hippo pathway regulates stem cells during homeostasis and regeneration of the flatworm Macrostomum lignano. Stem Cells Dev *22*, 2174-185.

Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell *2*, 819-830.

Denker, E., Manuel, M., Leclère, L., Le Guyader, H., and Rabet, N. (2008). Ordered progression of nematogenesis from stem cells through differentiation stages in the tentacle bulb of Clytia hemisphaerica (Hydrozoa, Cnidaria). Dev Biol *315*, 99-113.

Doe, C.Q., Fuerstenberg, S., and Peng, C.Y. (1998). Neural stem cells: from fly to vertebrates. J Neurobiol *36*, 111-127.

Edgecombe, G.D., Giribet, G., Dunn, C.W., Hejnol, A., Kristensen, R.M., Neves, R.C., Rouse, G.W., Worsaae, K., and Sørensen, M.V. (2011). Higher-level metazoan relationships: recent progress and remaining questions. Organisms Diversity & Evolution *11*, 151-172.

Egger, B., Gschwentner, R., and Rieger, R. (2007). Free-living flatworms under the knife: past and present. Dev Genes Evol *217*, 89-104.

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

Egger, B., Steinke, D., Tarui, H., De Mulder, K., Arendt, D., Borgonie, G., Funayama, N., Gschwentner, R., Hartenstein, V., et al. (2009). To be or not to be a flatworm: the acoel controversy. PLoS One *4*, e5502.

Eisenhoffer, G.T., Kang, H., and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian Schmidtea mediterranea. Cell Stem Cell *3*, 327-339.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-56.

Extavour, C.G., and Akam, M. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. Development *130*, 5869-884.

Felix, D.A., and Aboobaker, A.A. (2010). The TALE class homeobox gene Smed-prep defines the anterior compartment for head regeneration. PLoS Genet *6*, e1000915.

Fernandéz-Taboada, E., Moritz, S., Zeuschner, D., Stehling, M., Schöler, H.R., Saló, E., and Gentile, L. (2010). Smed-SmB, a member of the LSm protein superfamily, is essential for chromatoid body organization and planarian stem cell proliferation. Development *137*, 1055-065.

Flachsova, M., Sindelka, R., and Kubista, M. (2013). Single blastomere expression profiling of Xenopus laevis embryos of 8 to 32-cells reveals developmental asymmetry. Sci Rep *3*, 2278.

Foley, J.E., Maeder, M.L., Pearlberg, J., Joung, J.K., Peterson, R.T., and Yeh, J.R. (2009). Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. Nat Protoc *4*, 1855-867.

Friedländer, M.R., Adamidi, C., Han, T., Lebedeva, S., Isenbarger, T.A., Hirst, M., Marra, M., Nusbaum, C., Lee, W.L., et al. (2009). High-resolution profiling and discovery of planarian small RNAs. Proc Natl Acad Sci U S A *106*, 11546-551.

Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M., and Noce, T. (1994). Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosophila vasa and its specific expression in germ cell lineage. Proc Natl Acad Sci U S A *91*, 12258-262.

Fulle, S., Centurione, L., Mancinelli, R., Sancilio, S., Manzoli, F.A., and Di Pietro, R. (2012). Stem cell ageing and apoptosis. Curr Pharm Des *18*, 1694-1717.

Fuller, M.T., and Spradling, A.C. (2007). Male and female Drosophila germline stem cells: two versions of immortality. Science *316*, 402-04.

Funayama, N. (2010). The stem cell system in demosponges: insights into the origin of somatic stem cells. Dev Growth Differ *52*, 1-14.

Funayama, N. (2013). The stem cell system in demosponges: suggested involvement of two types of cells: archeocytes (active stem cells) and choanocytes (food-entrapping flagellated cells). Dev Genes Evol *223*, 23-38.

Funayama, N., Nakatsukasa, M., Mohri, K., Masuda, Y., and Agata, K. (2010). Piwi expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. Evol Dev *12*, 275-287.

Galliot, B. (2012). Hydra, a fruitful model system for 270 years. Int J Dev Biol *56*, 411-423.

Galliot, B. (2013). Injury-induced asymmetric cell death as a driving force for head regeneration in Hydra. Dev Genes Evol *223*, 39-52.

Gargioli, C., and Slack, J.M. (2004). Cell lineage tracing during Xenopus tail regeneration. Development *131*, 2669-679.

Gentile, L., Cebrià, F., and Bartscherer, K. (2011). The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration. Dis Model Mech *4*, 12-19.

González-Estévez, C., and Saló, E. (2010). Autophagy and apoptosis in planarians. Apoptosis *15*, 279-292.

Grskovic, M., Javaherian, A., Strulovici, B., and Daley, G.Q. (2011). Induced pluripotent stem cells-

-opportunities for disease modelling and drug discovery. Nat Rev Drug Discov *10*, 915-929.

Gschwentner, R., Ladurner, P., Nimeth, K., and Rieger, R. (2001). Stem cells in a basal bilaterian. Cell Tissue Res *304*, 401-08.

Guo, T., Peters, A.H., and Newmark, P.A. (2006). A Bruno-like gene is required for stem cell maintenance in planarians. Dev Cell *11*, 159-169.

Gurdon, J.B. (2006). From nuclear transfer to nuclear reprogramming: the reversal of cell differentiation. Annu Rev Cell Dev Biol *22*, 1-22.

Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008). Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science *319*, 323-27.

Handberg-Thorsager, M., Fernandez, E., and Salo, E. (2007). Stem cells and regeneration in planarians. Frontiers in bioscience: a journal and virtual library *13*, 6374-394.

Hanna, J., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., Creyghton, M.P., Steine, E.J., Cassady, J.P., et al. (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. Cell *133*, 250-264.

Hartwig, F.P., and Collares, T. (2013). Telomere dysfunction and tumor suppression responses in dyskeratosis congenita: balancing cancer and tissue renewal impairment. Ageing Res Rev *12*, 642-652.

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

Hay, E.D., and Fischman, D.A. (1961). Origin of the blastema in regenerating limbs of the newt Triturus viridescens. An autoradiographic study using tritiated thymidine to follow cell proliferation and migration. Dev Biol *3*, 26-59.

Hayashi, T., and Agata, K. (2012). A unique FACS method to isolate stem cells in planarian. Methods Mol Biol *879*, 29-37.

Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Dev Growth Differ *48*, 371-380.

Heasman, J. (1997). Patterning the Xenopus blastula. Development *124*, 4179-191.

Hemmrich, G., and Bosch, T.C. (2008). Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. Bioessays *30*, 1010-18.

Hernández, J., Torres-Espín, A., and Navarro, X. (2011). Adult stem cell transplants for spinal cord injury repair: current state in preclinical research. Curr Stem Cell Res Ther *6*, 273-287.

Higuchi, S., Hayashi, T., Hori, I., Shibata, N., Sakamoto, H., and Agata, K. (2007). Characterization and categorization of fluorescence activated cell sorted planarian stem cells by ultrastructural analysis. Dev Growth Differ *49*, 571-581.

Hodgkinson, T., Yuan, X.F., and Bayat, A. (2009). Adult stem cells in tissue engineering. Expert Rev Med Devices *6*, 621-640.

Holland, J.D., Klaus, A., Garratt, A.N., and Birchmeier, W. (2013). Wnt signaling in stem and cancer stem cells. Curr Opin Cell Biol *25*, 254-264.

Hombach-Klonisch, S., Panigrahi, S., Rashedi, I., Seifert, A., Alberti, E., Pocar, P., Kurpisz, M., Schulze-Osthoff, K., Mackiewicz, A., and Los, M. (2008). Adult stem cells and their trans-differentiation potential--perspectives and therapeutic applications. J Mol Med (Berl) *86*, 1301-314.

Hori, I. (1997). Cytological approach to morphogenesis in the planarian blastema. II. The effect of neuropeptides. J Submicrosc Cytol Pathol *29*, 91-97.

Houghton, J., Morozov, A., Smirnova, I., and Wang, T.C. (2007). Stem cells and cancer. Semin Cancer Biol *17*, 191-203.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D.V., Blaser, H., Raz, E., et al. (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69-82.

James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273-282.

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D.B., and Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution *67*, 3233-242.

Johnson, S.L., and Weston, J.A. (1995). Temperature-sensitive mutations that cause stage-specific defects in Zebrafish fin regeneration. Genetics *141*, 1583-595.

Jones, C.M., and Smith, J.C. (2008). An overview of Xenopus development. Methods Mol Biol *461*, 385-394.

Juliano, C., Wang, J., and Lin, H. (2011). Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. Annu Rev Genet *45*, 447-469.

Kanno, H. (2013). Regenerative therapy for neuronal diseases with transplantation of somatic stem cells. World J Stem Cells *5*, 163-171.

Kawamura, K., and Sunanaga, T. (2010). Hemoblasts in colonial tunicates: are they stem cells or tissue-restricted progenitor cells? Dev Growth Differ *52*, 69-76.

Kawamura, K., Sugino, Y., Sunanaga, T., and Fujiwara, S. (2008). Multipotent epithelial cells in the process of regeneration and asexual reproduction in colonial tunicates. Dev Growth Differ *50*, 1-11.

Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322*, 1065-69.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., et al. (2010). Epigenetic memory in induced pluripotent stem cells. Nature *467*, 285-290.

Kim, P.T., and Ong, C.J. (2012). Differentiation of definitive endoderm from mouse embryonic stem cells. Results Probl Cell Differ *55*, 303-319.

Kirilly, D., and Xie, T. (2007). The Drosophila ovary: an active stem cell community. Cell Res *17*, 15-25.

Knoepfler, P.S. (2009). Deconstructing stem cell tumorigenicity: a roadmap to safe regenerative medicine. Stem Cells *27*, 1050-56.

Komiya, T., Itoh, K., Ikenishi, K., and Furusawa, M. (1994). Isolation and characterization of a novel gene of the DEAD box protein family which is specifically expressed in germ cells of Xenopus laevis. Dev Biol *162*, 354-363.

Krause, D.S., Scadden, D.T., and Preffer, F.I. (2013). The hematopoietic stem cell niche--home for friend and foe? Cytometry B Clin Cytom *84*, 7-20.

Krause, D.S., Theise, N.D., Collector, M.I., Henegariu, O., Hwang, S., Gardner, R., Neutzel, S., and Sharkis, S.J. (2001). Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. Cell *105*, 369-377.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., et al. (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839-849.

Labbé, R.M., Irimia, M., Currie, K.W., Lin, A., Zhu, S.J., Brown, D.D., Ross, E.J., Voisin, V., Bader, G.D., et al. (2012). A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. Stem Cells *30*, 1734-745.

Ladurner, P., Pfister, D., Seifarth, C., Schärer, L., Mahlknecht, M., Salvenmoser, W., Gerth, R., Marx, F., and Rieger, R. (2005). Production and characterisation of cell- and tissue-specific monoclonal antibodies for the flatworm Macrostomum sp. Histochem Cell Biol *123*, 89-104.

Ladurner, P., Rieger, R., and Baguñà, J. (2000). Spatial distribution and differentiation potential of stem cells in hatchlings and adults in the marine platyhelminth macrostomum sp.: a bromodeoxyuridine analysis. Dev Biol *226*, 231-241.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Laird, D.J., De Tomaso, A.W., and Weissman, I.L. (2005). Stem cells are units of natural selection in a colonial ascidian. Cell *123*, 1351-360.

Lasko, P.F., and Ashburner, M. (1988). The product of the Drosophila gene vasa is very similar to eukaryotic initiation factor-4A. Nature *335*, 611-17.

Lepilina, A., Coon, A.N., Kikuchi, K., Holdway, J.E., Roberts, R.W., Burns, C.G., and Poss, K.D. (2006). A dynamic epicardial injury response supports progenitor cell activity during zebrafish heart regeneration. Cell *127*, 607-619.

Li, Z., and Chen, Y.G. (2013). Functions of BMP signaling in embryonic stem cell fate determination. Exp Cell Res *319*, 113-19.

Li, Z., Fei, T., Zhang, J., Zhu, G., Wang, L., Lu, D., Chi, X., Teng, Y., Hou, N., et al. (2012). BMP4 Signaling Acts via dual-specificity phosphatase 9 to control ERK activity in mouse embryonic stem cells. Cell Stem Cell *10*, 171-182.

Lian, I., Kim, J., Okazawa, H., Zhao, J., Zhao, B., Yu, J., Chinnaiyan, A., Israel, M.A., Goldstein, L.S., et al. (2010). The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation. Genes Dev *24*, 1106-118.

Lien, C.-L., Schebesta, M., Makino, S., Weber, G.J., and Keating, M.T. (2006). Gene expression analysis of zebrafish heart regeneration. PLoS Biol *4*, e260.

Lin, G., Chen, Y., and Slack, J.M. (2007). Regeneration of neural crest derivatives in the Xenopus tadpole tail. BMC Dev Biol *7*, 56.

Lin, H.T., Otsu, M., and Nakauchi, H. (2013). Stem cell therapy: an exercise in patience and prudence. Philos Trans R Soc Lond B Biol Sci *368*, 20110334.

Liu, S.Y., Selck, C., Friedrich, B., Lutz, R., Vila-Farré, M., Dahl, A., Brandl, H., Lakshmanaperumal, N., Henry, I., and Rink, J.C. (2013). Reactivating head regrowth in a regeneration-deficient planarian species. Nature *500*, 81-84.

Loh, Y.H., Agarwal, S., Park, I.H., Urbach, A., Huo, H., Heffner, G.C., Kim, K., Miller, J.D., Ng, K., and Daley, G.Q. (2009). Generation of induced pluripotent stem cells from human blood. Blood *113*, 5476-79.

Losick, V.P., Morris, L.X., Fox, D.T., and Spradling, A. (2011). Drosophila stem cell niches: a decade of discovery suggests a unified view of stem cell regulation. Dev Cell *21*, 159-171.

Macarthur, B.D., Ma'ayan, A., and Lemischka, I.R. (2009). Systems biology of stem cell fate and cellular reprogramming. Nat Rev Mol Cell Biol *10*, 672-681.

Margolis, J., and Spradling, A. (1995). Identification and behavior of epithelial stem cells in the Drosophila ovary. Development *121*, 3797-3807.

Marión, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149-153.

Martinez, E.C., and Kofidis, T. (2011). Adult stem cells for cardiac tissue engineering. J Mol Cell Cardiol *50*, 312-19.

Martinez-Agosto, J.A., Mikkola, H.K., Hartenstein, V., and Banerjee, U. (2007). The hematopoietic stem cell and its niche: a comparative view. Genes Dev *21*, 3044-060.

McCammon, J.M., Doyon, Y., and Amacher, S.L. (2011). Inducing high rates of targeted mutagenesis in zebrafish using zinc finger nucleases (ZFNs). Methods Mol Biol *770*, 505-527.

McDougall, A., Chenevert, J., Lee, K.W., Hebras, C., and Dumollard, R. (2011). Cell cycle in ascidian eggs and embryos. Results Probl Cell Differ *53*, 153-169.

Micchelli, C.A., and Perrimon, N. (2006). Evidence that stem cells reside in the adult Drosophila midgut epithelium. Nature *439*, 475-79.

Mimeault, M., and Batra, S.K. (2012). Great promise of tissue-resident adult stem/progenitor cells in transplantation and cancer therapies. Adv Exp Med Biol *741*, 171-186.

Mitalipov, S., and Wolf, D. (2009). Totipotency, pluripotency and nuclear reprogramming. Adv Biochem Eng Biotechnol *114*, 185-199.

Moad, M., Pal, D., Hepburn, A.C., Williamson, S.C., Wilson, L., Lako, M., Armstrong, L., Hayward, S.W., Franco, O.E., et al. (2013). A novel model of urinary tract differentiation, tissue regeneration, and disease: reprogramming human prostate and bladder cells into induced pluripotent stem cells. Eur Urol *64*, 753-761.

Mochizuki, K., Nishimiya-Fujisawa, C., and Fujisawa, T. (2001). Universal occurrence of the vasa-related genes among metazoans and their germline expression in Hydra. Dev Genes Evol *211*, 299-308.

Mochizuki, K., Sano, H., Kobayashi, S., Nishimiya-Fujisawa, C., and Fujisawa, T. (2000). Expression and evolutionary conservation of nanos-related genes in Hydra. Dev Genes Evol *210*, 591-602.

Moon, S.H., Kim, J.M., Hong, K.S., Shin, J.M., Kim, J., and Chung, H.M. (2011). Differentiation of hESCs into Mesodermal Subtypes: Vascular-, Hematopoietic- and Mesenchymal-lineage Cells. Int J Stem Cells *4*, 24-34.

Morgan, T.H. (1898). Experimental Studies of the Regeneration of Planaria Maculata. Arch. Entwicklungsmech. Org. *7*, 364-397.

Morgan, T.H. (1902). Growth and regeneration in Planaria lugubris. Arch. Ent. Mech. Org. *13*

*M***orita, M.,** and Best, J.B. (1984). Electron microscopic studies of planarian regeneration. III. Degeneration and differentiation of muscles. Journal of Experimental Zoology *229*, 413-424.

Morris, J., Ladurner, P., Rieger, R., Pfister, D., Del Mar De Miguel-Bonet, M., Jacobs, D., and Hartenstein, V. (2006). The Macrostomum lignano EST database as a molecular resource for studying platyhelminth development and phylogeny. Dev Genes Evol *216*, 695-707.

Morris, J., Nallur, R., Ladurner, P., Egger, B., Rieger, R., and Hartenstein, V. (2004). The embryonic development of the flatworm Macrostomum sp. Dev Genes Evol *214*, 220-239.

Morrison, J.I., Lööf, S., He, P., and Simon, A. (2006). Salamander limb regeneration involves the activation of a multipotent skeletal muscle satellite cell population. J Cell Biol *172*, 433-440.

Morrison, S.J., and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. Nature *441*, 1068-074.

Morrison, S.J., and Spradling, A.C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. Cell *132*, 598-611.

Mouton, S., Willems, M., Braeckman, B.P., Egger, B., Ladurner, P., Schärer, L., and Borgonie, G. (2009). The free-living flatworm Macrostomum lignano: a new model organism for ageing research. Exp Gerontol *44*, 243-49.

De Mulder, K., Kuales, G., Pfister, D., Egger, B., Seppi, T., Eichberger, P., Borgonie, G., and Ladurner, P. (2010). Potential of Macrostomum lignano to recover from gamma-ray irradiation. Cell Tissue Res *339*, 527-542.

De Mulder, K., Kuales, G., Pfister, D., Willems, M., Egger, B., Salvenmoser, W., Thaler, M., Gorny, A.K., Hrouda, M., et al. (2009). Characterization of the stem cell system of the acoel Isodiametra pulchra. BMC Dev Biol *9*, 69.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., et al. (2009). Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. Dev Biol *334*, 198-212.

Mullen, A.C., Orlando, D.A., Newman, J.J., Lovén, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-β signaling. Cell *147*, 565-576.

Müller, W.A., Teo, R., and Frank, U. (2004). Totipotent migratory stem cells in a hydroid. Dev Biol *275*, 215-224.

Müller, W.E. (2006). The stem cell concept in sponges (Porifera): Metazoan traits. Semin Cell Dev Biol *17*, 481-491.

Nakauchi, M. (1982). Asexual development of ascidians: its biological significance, diversity, and morphogenesis. American Zoologist *22*, 753-763.

Nelson, C.M., Ackerman, K.M., O'Hayer, P., Bailey, T.J., Gorsuch, R.A., and Hyde, D.R. (2013). Tumor Necrosis Factor-Alpha Is Produced by Dying Retinal Neurons and Is Required for Müller Glia Proliferation during Zebrafish Retinal Regeneration. The Journal of Neuroscience *33*, 6524-539.

Neumüller, R.A., and Knoblich, J.A. (2009). Dividing cellular asymmetry: asymmetric cell division and its implications for stem cells and cancer. Genes Dev *23*, 2675-699.

Newmark, P.A., and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev Biol *220*, 142-153.

Newmark, P.A., and Sánchez Alvarado, A. (2002). Not your father's planarian: a classic model enters the era of functional genomics. Nat Rev Genet *3*, 210-19.

Newmark, P.A., Reddien, P.W., Cebrià, F., and Sánchez Alvarado, A. (2003). Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. Proc Natl Acad Sci U S A *100 Suppl 1*, 11861-65.

Nielen, M.G., de Vries, S.A., and Geijsen, N. (2013). European stem cell research in legal shackles. EMBO J

Nimeth, K., Ladurner, P., Gschwentner, R., Salvenmoser, W., and Rieger, R. (2002). Cell renewal and apoptosis in Macrostomum sp.[Lignano]. Cell biology international *26*, 801-815.

Nimeth, K.T., Mahlknecht, M., Mezzanato, A., Peter, R., Rieger, R., and Ladurner, P. (2004). Stem cell dynamics during growth, feeding, and starvation in the basal flatworm Macrostomum sp. (Platyhelminthes). Dev Dyn *230*, 91-99.

Nishimura, O., Hirao, Y., Tarui, H., and Agata, K. (2012). Comparative transcriptome analysis between planarian Dugesia japonica and other platyhelminth species. BMC Genomics *13*, 289.

Noggle, S.A., Weiler, D., and Condie, B.G. (2006). Notch signaling is inactive but inducible in human embryonic stem cells. Stem Cells *24*, 1646-653.

Nüsslein-Volhard, C., Frohnhöfer, H.G., and Lehmann, R. (1987). Determination of anteroposterior polarity in Drosophila. Science *238*, 1675-681.

Ohlstein, B., and Spradling, A. (2006). The adult Drosophila posterior midgut is maintained by pluripotent stem cells. Nature *439*, 470-74.

Ohlstein, B., and Spradling, A. (2007). Multipotent Drosophila intestinal stem cells specify daughter cell fates by differential notch signaling. Science *315*, 988-992.

Okano, H., Nakamura, M., Yoshida, K., Okada, Y., Tsuji, O., Nori, S., Ikeda, E., Yamanaka, S., and Miura, K. (2013). Steps toward safe cell therapy using induced pluripotent stem cells. Circ Res *112*, 523-533.

Onal, P., Grün, D., Adamidi, C., Rybak, A., Solana, J., Mastrobuoni, G., Wang, Y., Rahn, H.P., Chen, W., et al. (2012). Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. EMBO J *31*, 2755-769.

Orii, H., Sakurai, T., and Watanabe, K. (2005). Distribution of the stem cells (neoblasts) in the planarian Dugesia japonica. Dev Genes Evol *215*, 143-157.

Oviedo, N.J., Nicolas, C.L., Adams, D.S., and Levin, M. (2008). Live Imaging of Planarian Membrane Potential Using DiBAC4(3). CSH Protoc *2008*, pdb.prot5055.

Pan, G., and Thomson, J.A. (2007). Nanog and transcriptional networks in embryonic stem cell pluripotency. Cell Res *17*, 42-49.

Paré, J.F., and Sherley, J.L. (2006). Biological principles for ex vivo adult stem cell expansion. Curr Top Dev Biol *73*, 141-171.

Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008). Disease-specific induced pluripotent stem cells. Cell *134*, 877-886.

Passier, R., and Mummery, C. (2003). Origin and use of embryonic and adult stem cells in differentiation and tissue repair. Cardiovasc Res *58*, 324-335.

Patricio, V.J. (2011). Science and Ethics: Bridge to the Future for Regenerative Medicine. Int J Stem Cells *4*, 79-84.

Pearl, J.I., Kean, L.S., Davis, M.M., and Wu, J.C. (2012). Pluripotent stem cells: immune to the immune system? Sci Transl Med *4*, 164ps25.

Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., and Sánchez Alvarado, A. (2009). Formaldehyde-based whole-mount in situ hybridization method for planarians. Dev Dyn *238*, 443-450.

Pearson, J., López-Onieva, L., Rojas-Ríos, P., and González-Reyes, A. (2009). Recent advances in Drosophila stem cell biology. Int J Dev Biol *53*, 1329-339.

Pedersen, K.J. (1959). Cytological studies on the planarian neoblast. Z Zellforsch Mikrosk Anat *50*, 799-817.

Pellettieri, J., and Sánchez Alvarado, A. (2007). Cell turnover and adult tissue homeostasis: from humans to planarians. Annu Rev Genet *41*, 83-105.

Petersen, C.P., and Reddien, P.W. (2008). Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. Science *319*, 327-330.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. (2007). The exceptional stem cell system of Macrostomum lignano: screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. Front Zool *4*, 9.

Philippe, H., Brinkmann, H., Copley, R.R., Moroz, L.L., Nakano, H., Poustka, A.J., Wallberg, A., Peterson, K.J., and Telford, M.J. (2011). Acoelomorph flatworms are deuterostomes related to Xenoturbella. Nature *470*, 255-58

Plickert, G., Frank, U., and Müller, W.A. (2012). Hydractinia, a pioneering model for stem cell biology and reprogramming somatic cells to pluripotency. Int J Dev Biol *56*, 519-534.

Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol *28*, 848-855.

Poss, K.D., Keating, M.T., and Nechiporuk, A. (2003). Tales of regeneration in zebrafish. Dev Dyn *226*, 202-210.

Poss, K.D., Nechiporuk, A., Hillam, A.M., Johnson, S.L., and Keating, M.T. (2002). Mps1 defines a proximal blastemal proliferative compartment essential for zebrafish fin regeneration. Development *129*, 5141-49.

Postlethwait, J.H., Woods, I.G., Ngo-Hazelett, P., Yan, Y.L., Kelly, P.D., Chu, F., Huang, H., Hill-Force, A., and Talbot, W.S. (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res *10*, 1890-1902.

Raff, M. (2003). Adult stem cell plasticity: fact or artifact? Annu Rev Cell Dev Biol *19*, 1-22.

Reddien, P.W., and Sánchez Alvarado, A. (2004). Fundamentals of planarian regeneration. Annu Rev Cell Dev Biol *20*, 725-757.

Reddien, P.W., Bermange, A.L., Murfitt, K.J., Jennings, J.R., and Sánchez Alvarado, A. (2005). Identification of genes needed for regeneration, stem cell function, and tissue homeostasis by systematic gene perturbation in planaria. Dev Cell *8*, 635-649.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science *310*, 1327-330.

Re'em, T., and Cohen, S. (2011). Microenvironment Design for Stem Cell Fate Determination. Adv Biochem Eng Biotechnol.

Rieger, R.M., Tyler, S., Smith, J.P.S., Rieger, G. (1991). Platyhelminthes: Turbellaria. In Microscopic Anatomy of Invertebrates *3*, 7-140. Harrisson, F.W., Bogitsch, B.J. (eds). New York: Wiley-Liss.

Rieger, R.M., Legniti, A., Ladurner, P., Reiter, D., Asch, E., Salvenmoser, W., Schurmann, W., and Peter, R. (1999). Ultrastructure of neoblasts in microturbellaria: significance for understanding stem cells in free-living Platyhelminthes. Invertebrate reproduction & development *35*, 127-140.

Rink, J.C. (2013). Stem cell systems and regeneration in planaria. Dev Genes Evol *223*, 67-84.

Rink, J.C., Gurley, K.A., Elliott, S.A., and Sánchez Alvarado, A. (2009). Planarian Hh signaling regulates regeneration polarity and links Hh pathway evolution to cilia. Science *326*, 1406-410.

Rinkevich, Y., Paz, G., Rinkevich, B., and Reshef, R. (2007). Systemic bud induction and retinoic acid signaling underlie whole body regeneration in the urochordate Botrylloides leachi. PLoS Biol *5*, e71.

Risitano, A.M., Maciejewski, J.P., Selleri, C., and Rotoli, B. (2007). Function and malfunction of hematopoietic stem cells in primary bone marrow failure syndromes. Curr Stem Cell Res Ther *2*, 39-52.

Robb, S.M., Ross, E., and Sánchez Alvarado, A. (2008). SmedGD: the Schmidtea mediterranea genome database. Nucleic Acids Res *36*, D599-D606.

Robinton, D.A., and Daley, G.Q. (2012). The promise of induced pluripotent stem cells in research and therapy. Nature *481*, 295-305.

Rossi, L., Salvetti, A., Lena, A., Batistoni, R., Deri, P., Pugliesi, C., Loreti, E., and Gremigni, V. (2006). DjPiwi-1, a member of the PAZ-Piwi gene family, defines a subpopulation of planarian stem cells. Dev Genes Evol *216*, 335-346.

Rossi, L., Salvetti, A., Marincola, F.M., Lena, A., Deri, P., Mannini, L., Batistoni, R., Wang, E., and Gremigni, V. (2007). Deciphering the molecular machinery of stem cells: a look at the neoblast gene expression profile. Genome Biol *8*, R62.

Rouhana, L., Shibata, N., Nishimura, O., and Agata, K. (2010). Different requirements for conserved post-transcriptional regulators in planarian regeneration and stem cell maintenance. Dev Biol *341*, 429-443.

Ryffel, G.U., Werdien, D., Turan, G., Gerhards, A., Goosses, S., and Senkel, S. (2003). Tagging muscle cell lineages in development and tail regeneration using Cre recombinase in transgenic Xenopus. Nucleic Acids Res *31*, e44.

Saló, E. (2006). The power of regeneration and the stem-cell kingdom: freshwater planarians (Platyhelminthes). Bioessays *28*, 546-559.

Salvetti, A., Rossi, L., Bonuccelli, L., Lena, A., Pugliesi, C., Rainaldi, G., Evangelista, M., and Gremigni, V. (2009). Adult stem cell plasticity: neoblast

repopulation in non-lethally irradiated planarians. Dev Biol *328*, 305-314.

Salvetti, A., Rossi, L., Deri, P., and Batistoni, R. (2000). An MCM2-related gene is expressed in proliferating cells of intact and regenerating planarians. Dev Dyn *218*, 603-614.

Samuel, J.K. (2011). Functional Recovery of Spinal Cord Injury Following Application of Intralesional Bone Marrow Mononuclear Cells Embedded in Polymer Scaffold-Two Year Follow-up in a Canine. Journal of Stem Cell Research & Therapy

Sandmann, T., Vogg, M.C., Owlarn, S., Boutros, M., and Bartscherer, K. (2011). The head-regeneration transcriptome of the planarian Schmidtea mediterranea. Genome Biol *12*, R76.

Sarras, M.P. (2012). Components, structure, biogenesis and function of the Hydra extracellular matrix in regeneration, pattern formation and cell differentiation. Int J Dev Biol *56*, 567-576.

Sato, K., Shibata, N., Orii, H., Amikura, R., Sakurai, T., Agata, K., Kobayashi, S., and Watanabe, K. (2006). Identification and origin of the germline stem cells as revealed by the expression of nanos-related gene in planarians. Dev Growth Differ *48*, 615-628.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63.

Satoh, N. (1994). Developmental biology of ascidians (Cambridge University Press Cambridge).

Sánchez Alvarado, A. (2004). Regeneration and the need for simpler model organisms. Philos Trans R Soc Lond B Biol Sci *359*, 759-763.

Sánchez Alvarado, A. (2006). Planarian regeneration: its end is its beginning. Cell *124*, 241-45.

Sánchez Alvarado, A. (2007). Stem cells and the Planarian *Schmidtea mediterranea*. Comptes rendus biologies *330*, 498-503.

Sánchez Alvarado, A., and Kang, H. (2005). Multicellularity, stem cells, and the neoblasts of the planarian Schmidtea mediterranea. Exp Cell Res *306*, 299-308.

Sánchez Alvarado, A., and Newmark, P.A. (1999). Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proc Natl Acad Sci U S A *96*, 5049-054.

Scimone, M.L., Meisel, J., and Reddien, P.W. (2010). The Mi-2-like Smed-CHD4 gene is required for stem cell differentiation in the planarian Schmidtea mediterranea. Development *137*, 1231-241.

Seipel, K., Yanze, N., and Schmid, V. (2004). The germ line and somatic stem cell gene Cniwi in the jellyfish Podocoryne carnea. Int J Dev Biol *48*, 1-7.

Selvaraj, V., Plane, J.M., Williams, A.J., and Deng, W. (2010). Switching cell fate: the remarkable rise of induced pluripotent stem cells and lineage reprogramming technologies. Trends Biotechnol *28*, 214-223.

Serakinci, N., and Keith, W.N. (2006). Therapeutic potential of adult stem cells. Eur J Cancer *42*, 1243-46.

Sharpless, N.E., and DePinho, R.A. (2007). How stem cells age and why this makes us grow old. Nat Rev Mol Cell Biol *8*, 703-713.

Shibata, N., Hayashi, T., Fukumura, R., Fujii, J., Kudome-Takamatsu, T., Nishimura, O., Sano, S., Son, F., Suzuki, N., et al. (2012). Comprehensive gene expression analyses in pluripotent stem cells of a planarian, Dugesia japonica. Int J Dev Biol *56*, 93-102.

Shibata, N., Umesono, Y., Orii, H., Sakurai, T., Watanabe, K., and Agata, K. (1999). Expression of vasa(vas)-related genes in germline cells and totipotent somatic stem cells of planarians. Dev Biol *206*, 73-87.

Siegrist, S.E., and Doe, C.Q. (2006). Extrinsic cues orient the cell division axis in Drosophila embryonic neuroblasts. Development *133*, 529-536.

Sikes, J.M., and Newmark, P.A. (2013). Restoration of anterior regeneration in a planarian with limited regenerative ability. Nature *500*, 77-80.

Silva, J., and Smith, A. (2008). Capturing pluripotency. Cell *132*, 532-36.

Simanov, D., Mellaart-Straver, I., Sormacheva, I., and Berezikov, E. (2012). The Flatworm Macrostomum lignano Is a Powerful Model Organism for Ion Channel and Stem Cell Research. Stem Cells Int *2012*, 167265.

Singh, S.R., Liu, W., and Hou, S.X. (2007). The adult Drosophila malpighian tubules are maintained by multipotent stem cells. Cell Stem Cell *1*, 191-203.

Slack, J.M. (2008). Origin of stem cells in organogenesis. Science *322*, 1498-1501.

Slack, J.M., Beck, C.W., Gargioli, C., and Christen, B. (2004). Cellular and molecular mechanisms of regeneration in Xenopus. Philos Trans R Soc Lond B Biol Sci *359*, 745-751.

Snippert, H.J., and Clevers, H. (2011). Tracking adult stem cells. EMBO Rep *12*, 113-122.

Solana, J., Kao, D., Mihaylova, Y., Jaber-Hijazi, F., Malla, S., Wilson, R., and Aboobaker, A. (2012). Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNA-seq, RNA interference and irradiation approach. Genome biology *13*, R19.

Spradling, A., Drummond-Barbosa, D., and Kai, T. (2001). Stem cells find their niche. Nature *414*, 98-104.

Stadtfeld, M., Brennand, K., and Hochedlinger, K. (2008). Reprogramming of pancreatic beta cells into induced pluripotent stem cells. Curr Biol *18*, 890-94.

Steindler, D.A. (2007). Stem cells, regenerative medicine, and animal models of disease. Ilar Journal *48*, 323-338.

Stocum, D.L. (2000). Transplantation and culture techniques for the analysis of urodele limb regeneration. Methods Mol Biol *136*, 71-78.

Straube, W.L., and Tanaka, E.M. (2006). Reversibility of the differentiated state: regeneration in amphibians. Artif Organs *30*, 743-755.

Subrammaniyan, R., Amalorpavanathan, J., Shankar, R., Rajkumar, M., Baskar, S., Manjunath, S.R., Senthilkumar, R., Murugan, P., Srinivasan, V.R., and Abraham, S. (2011). Application of autologous bone marrow mononuclear cells in six patients with advanced chronic critical limb ischemia as a result of diabetes: our experience. Cytotherapy *13*, 993-99.

Sugawara, T., Nishino, K., Umezawa, A., and Akutsu, H. (2012). Investigating cellular identity and manipulating cell fate using induced pluripotent stem cells. Stem Cell Res Ther *3*, 8.

Sugino, Y.M., Matsumura, M., and Kawamura, K. (2007). Body muscle-cell differentiation from coelomic stem cells in colonial tunicates. Zoolog Sci *24*, 542-46.

Sugio, M., Takeuchi, K., Kutsuna, J., Tadokoro, R., Takahashi, Y., Yoshida-Noro, C., and Tochinai, S. (2008). Exploration of embryonic origins of germline stem cells and neoblasts in Enchytraeus japonensis (Oligochaeta, Annelida). Gene Expr Patterns *8*, 227-236.

Sugio, M., Yoshida-Noro, C., Ozawa, K., and Tochinai, S. (2012). Stem cells in asexual reproduction of Enchytraeus japonensis (Oligochaeta, Annelid): proliferation and migration of neoblasts. Dev Growth Differ *54*, 439-450.

Sui, L., Bouwens, L., and Mfopou, J.K. (2013). Signaling pathways during maintenance and definitive endoderm differentiation of embryonic stem cells. Int J Dev Biol *57*, 1-12.

Sunanaga, T., Saito, Y., and Kawamura, K. (2006). Postembryonic epigenesis of Vasa-positive germ cells from aggregated hemoblasts in the colonial ascidian, Botryllus primigenus. Dev Growth Differ *48*, 87-100.

Tadokoro, R., Sugio, M., Kutsuna, J., Tochinai, S., and Takahashi, Y. (2006). Early segregation of germ and somatic lineages during gonadal regeneration in the annelid Enchytraeus japonensis. Curr Biol *16*, 1012-17.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Takashima, S., Mkrtchyan, M., Younossi-Hartenstein, A., Merriam, J.R., and Hartenstein, V. (2008). The behaviour of Drosophila adult hindgut stem cells is controlled by Wnt and Hh signalling. Nature *454*, 651-55.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.R., Ueno, Y., Zheng, Y.W., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature *499*, 481-84.

Tamm, C., Böwer, N., and Annerén, C. (2011). Regulation of mouse embryonic stem cell self-renewal by a Yes-YAP-TEAD2 signaling pathway downstream of LIF. J Cell Sci *124*, 1136-144.

Tanaka, E.M. (2003). Regeneration: if they can do it, why can't we? Cell *113*, 559-562.

Terai, S., Ishikawa, T., Omori, K., Aoyama, K., Marumoto, Y., Urata, Y., Yokoyama, Y., Uchida, K., Yamasaki, T., et al. (2006). Improved liver function in patients with liver cirrhosis after autologous bone marrow cell infusion therapy. Stem Cells *24*, 2292-98.

Tiozzo, S., Brown, F.D., De Tomaso, A.W. (2008) Regeneration and stem cells in ascidians. In Stem Cells, 95-112. Springer.

Trembley, A. (1744). *Mémoires pour servir à l'histoire d'un genre de polypes d'eau douce, à bras en forme de cornes* (Leiden: Jean & Herman Verbeek).

Tsai, R.Y., Kittappa, R., and McKay, R.D. (2002). Plasticity, niches, and the use of stem cells. Dev Cell *2*, 707-712.

Umesono, Y., Tasaki, J., Nishimura, Y., Hrouda, M., Kawaguchi, E., Yazawa, S., Nishimura, O., Hosoda, K., Inoue, T., and Agata, K. (2013). The molecular logic for planarian regeneration along the anterior-posterior axis. Nature *500*, 73-76.

Umesono, Y., Watanabe, K., and Agata, K. (1997). A planarian orthopedia homolog is specifically expressed in the branch region of both the mature and regenerating brain. Dev Growth Differ *39*, 723-27.

Utikal, J., Maherali, N., Kulalert, W., and Hochedlinger, K. (2009). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. J Cell Sci *122*, 3502-510.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci *118*, 4495-4509.

Vanuytsel, T., Senger, S., Fasano, A., and Shea-Donohue, T. (2013). Major signaling pathways in intestinal stem cells. Biochim Biophys Acta *1830*, 2410-426.

Verdoodt, F., Bert, W., Couvreur, M., De Mulder, K., and Willems, M. (2012). Proliferative response of the stem cell system during regeneration of the rostrum in Macrostomum lignano (Platyhelminthes). Cell Tissue Res *347*, 397-406.

Vesterlund, L., Jiao, H., Unneberg, P., Hovatta, O., and Kere, J. (2011). The zebrafish transcriptome during early development. BMC Dev Biol *11*, 30.

Votteler, M., Kluger, P.J., Walles, H., and Schenke-Layland, K. (2010). Stem cell microenvironments--unveiling the secret of how stem cell fate is defined. Macromol Biosci *10*, 1302-315.

Wagner, D.E., Ho, J.J., and Reddien, P.W. (2012). Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. Cell Stem Cell *10*, 299-311.

Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. Science *332*, 811-16.

Walker, M.R., Patel, K.K., and Stappenbeck, T.S. (2009). The stem cell niche. J Pathol *217*, 169-180.

Wills, A.A., Holdway, J.E., Major, R.J., and Poss, K.D. (2008). Regulated addition of new myocardial and epicardial cells fosters homeostatic cardiac growth and maintenance in adult zebrafish. Development *135*, 183-192.

Wolff, E. (1948). Sur la migration des cellules de régénération chez les planaires. Rev. Suisse Zool. *55*, 218-227.

Woods, I.G., Kelly, P.D., Chu, F., Ngo-Hazelett, P., Yan, Y.-L., Huang, H., Postlethwait, J.H., and Talbot, W.S. (2000). A comparative map of the zebrafish genome. Genome Res *10*, 1903-914.

Wu, D.C., Boyd, A.S., and Wood, K.J. (2007). Embryonic stem cell transplantation: potential applicability in cell replacement therapy and regenerative medicine. Front Biosci *12*, 4525-535.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281-292.

Yoon, C., Kawakami, K., and Hopkins, N. (1997). Zebrafish vasa homologue RNA is localized to the cleavage planes of 2-and 4-cell-stage embryos and is expressed in the primordial germ cells. Development *124*, 3157-165.

Yoshida-Noro, C., and Tochinai, S. (2010). Stem cell system in asexual and sexual reproduction of Enchytraeus japonensis (Oligochaeta, Annelida). Dev Growth Differ *52*, 43-55.

Yu, F., Kuo, C.T., and Jan, Y.N. (2006). Drosophila neuroblast asymmetric cell division: recent advances and implications for stem cell biology. Neuron *51*, 13-20.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-920.

Zhan, M., Miura, T., Xu, X., and Rao, M.S. (2005). Conservation and variation of gene regulation in embryonic stem cells assessed by comparative genomics. Cell Biochem Biophys *43*, 379-405.

Zhao, T., Zhang, Z.N., Rong, Z., and Xu, Y. (2011). Immunogenicity of induced pluripotent stem cells. Nature *474*, 212-15.

Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., et al. (2009). Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell *4*, 381-84.

CHAPTER 2

INITIAL DRAFT GENOME ASSEMBLIES OF THE FLATWORMS *MACROSTOMUM LIGNANO* AND *MACROSTOMUM HYSTRIX* AND THE ACOEL *ISODIAMETRA PULCHRA*

Daniil Simanov[1], Kaja Wasik[2], Andres Canela[2], Stijn Mouton[3], Patrick van Zon[4], Katrien de Mulder[1], Ewart de Bruijn[1], Edwin Cuppen[1], Aziz Aboobaker[5], Weihong Qi[6], Andreas Hejnol[7], Dita B Vizoso[8], Steven Ramm[9], Lukas Schärer[8], Michael Schatz[2], Gregory J Hannon[2], Peter Ladurner[10] and Eugene Berezikov[1,3].

[1]Hubrecht Institute, Hubrecht Institute, KNAW and University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands
[2]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
[3]European Research Institute for the Biology of Ageing, University of Groningen, University Medical Centre Groningen, 9713AV Groningen, The Netherlands
[4]University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands
[5]University of Oxford, OX1 3PS Oxford, United Kingdom
[6]Functional Genomics Center Zurich, 8057 Zurich, Switzerland
[7]Sars International Centre for Marine Molecular Biology, N-5008 Bergen, Norway8University of Basel, Zoological Institute, Basel, Switzerland
[9]Bielefeld University, 33615 Bielefeld, Germany
[10]University of Innsbruck, Innsbruck, Austria

Work in progress.

ABSTRACT

Flatworms (phylum Platyhelminthes) are relatively simple bilaterian invertebrates characterized by unsegmented body, lack of body cavity and specialized circulatory and respiratory organs. They represent the oldest clade of bilaterian animals, and the sub-taxon Macrostomorpha is the most basal group of the flatworms. A member of this group, *Macrostomum lignano*, has recently emerged as a model organism for studies of stem cell biology, regeneration, sex allocation and ageing. Another group of seemingly primitive soft-bodied worms known as Acoels was until recently considered to be the oldest members of Platyhelminthes. However, recently presented molecular and developmental data makes the Acoela a completely separate group with a debated phylogenetic position either at the very base of Bilateria or much closer to chordates. Here we report first results towards sequencing and assembly of genomes of *M. lignano*, its close relative *Macrostomum hystrix* and acoel *Isodiametra pulchra* and argue how comparison of these genomes might shed light on the evolution of adult stem cell pluripotency in bilaterian species.

INTRODUCTION

Macrostomum lignano (Fig. 1 D-F) is a member of the order Macrostomida inside the not clearly defined sub-taxon Macrostomorpha of the phylum Platyhelminthes (Ladurner et al., 2005). Platyhelminthes, also known as flatworms, are unsegmented soft-bodied animals with relatively simple morphology, lacking body cavity and specialized circulatory and respiratory organ systems. Partly because of these last two characteristics flatworms are regarded as a primitive stage in the evolution of bilaterians, and members of Macrostomorpha are, in turn, among the most basal Platyhelminthes (Ehlers, 1985; Ehlers and Sopott-Ehlers, 1995; Riutort et al., 2012). Most of them are small-sized (up to 5 mm) highly transparent animals that inhabit freshwater, brackish and marine environments. Species *M. lignano*, collected on the Italian coast of the Northern Adriatic, is successfully cultured in the laboratory conditions for more than a decade (Ladurner et al., 2000; Ladurner et al., 2005).

Another group of basal animals known as Acoels was traditionally considered to belong to Platyhelminthes. In 2004 new molecular and developmental approaches demonstrated that Acoelomorpha represents a separate phylum (Baguñà and Riutort, 2004), although its position in the tree of life is still under debate. Some researchers believe Acoels to be a group at the very base of Bilateria, even older than Platyhelminthes and just slightly more derived than cnidarians (Egger et al., 2009; Edgecombe et al., 2011). Others say that they may lie near the base of the deuterostomes and, thus, much closer to chordates (Philippe et al., 2011; Maxmen, 2011). Acoels have even simpler anatomy than flatworms, with no circulatory, respiratory or even excretory systems and no true gonads. They are almost entirely marine, living on the algae, in the sediment or swimming as plankton. *Isodiametra pulchra* (Fig. 1 A-C) is one of the best-studied animals within this group that can be grown and maintained in the laboratory conditions (De Mulder et al., 2009).

Acoels and flatworms are the only two groups of animals with bilateral symmetry that possess unique population of adult pluripotent stem cells, also called neoblasts (Fig. 1 C and F). These cells are responsible for impressive regeneration abilities and tissue homeostasis in the whole organism (Gschwentner et al., 2001; De Mulder et al., 2009; Baguñà, 2012). There are highly similar stem cell populations described for sponges and cnidarians (Bode, 1996; Funayama, 2010; Müller, 2006). Just as neoblasts of flatworms and acoels, adult stem cells (known as interstitial cells) of cnidarians *Hydra* and *Nematostella vectensis*, are able to differentiate into every cell type (Darling et al., 2005; Watanabe et al., 2009; Bosch et al., 2010; Galliot, 2012). The same applies to archaeocytes of sponges (Funayama, 2010; Funayama, 2013). It is not clear whether these stem cell populations are evolutionary conserved, and how the pluripotency is established and maintained in all four phyla. More generally, we do not know whether same molecular and physiological mechanisms are involved in pluripotent stem cell functioning in bilaterians and more ancient animals with radial symmetry. Evolutionary conservation of the stem cell regulation has only been studied for a little more than a decade (Benfey, 1999; Zhan et al., 2005), and today we only have first indications that early branching metazoans possess many if not most of stem cell specific genes characteristic for vertebrates, even mammals (Bosch, 2008; Hemmrich and Bosch, 2008). The genomes of *Hydra magnipapillata (Chapman et al., 2010)*, *N. vectensis* (Putnam et al., 2007) and the sponge *Amphimedon queenslandica* (Srivastava et al., 2010) were

sequenced recently, and the comparison of the gene sets of these animals with the data from flatworms and acoels would give a new valuable insight into the evolution of pluripotency.

M. lignano is currently used as a model organism in a number of laboratories and recently has proven itself useful for stem cell, regeneration, aging and sex allocation research (Egger et

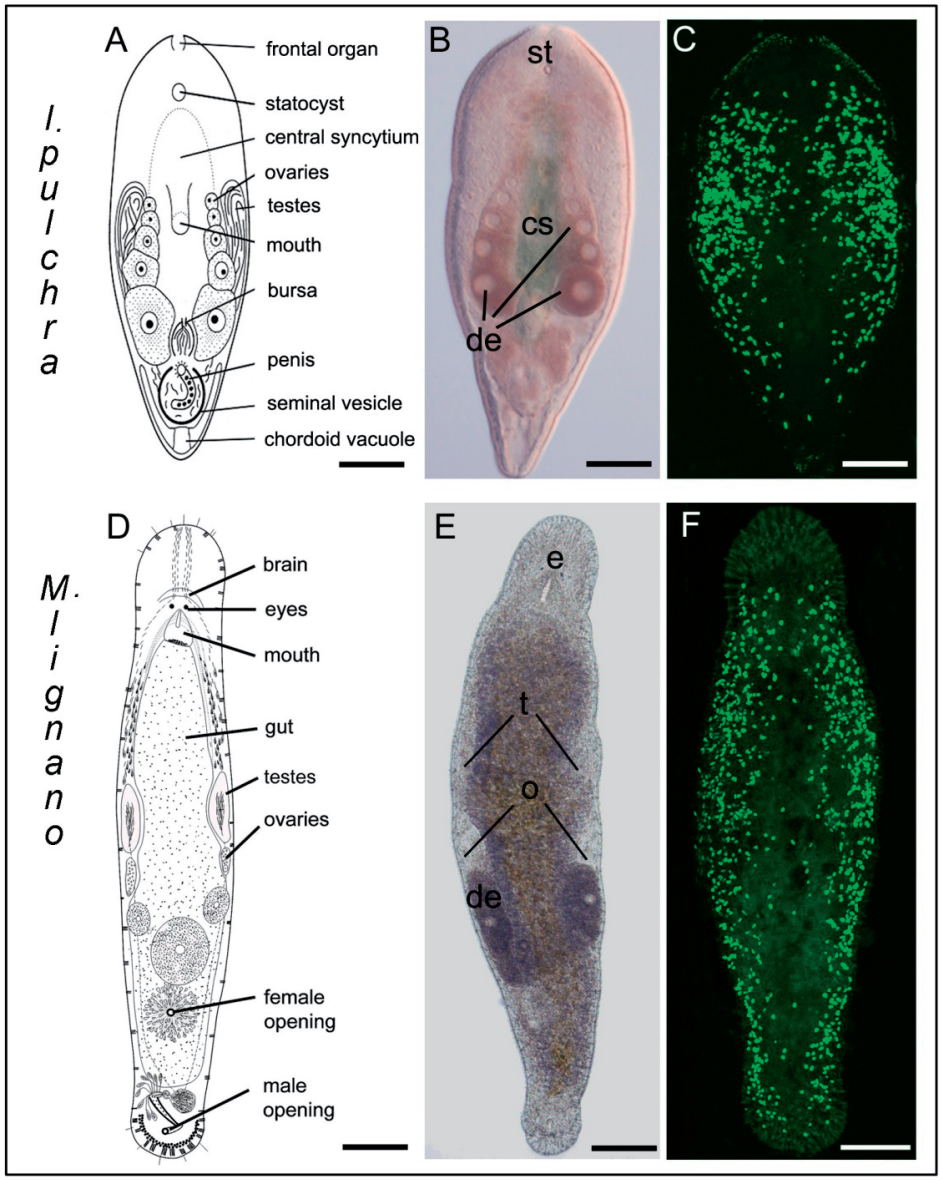


Figure 1. *M. lignano* and *I. pulchra*. **(A, D)** Schematic overview. **(B, F)** DIC-image of living specimens. **(C, F)** Distribution of proliferating cells (neoblasts), as shown by BrdU labelling.(t) testes, (o) ovaries, (de) developing eggs, (st) statocyst, (cs) central syncytium, (e) eyes. Scale bars 100 µm. From (De Mulder, 2009).

al., 2006; Pfister et al., 2008; De Mulder et al., 2009; Mouton et al., 2009; Verdoodt et al., 2012; Simanov et al., 2012; Janicke et al., 2013; Demircan and Berezikov, 2013). The neoblast system of *M. lignano* has been described in a fair detail (Rieger et al., 1999; Bode et al., 2006; De Mulder et al., 2009), and an impressive toolkit has been developed for this model organism over the last 8 years (Ladurner et al., 2005; Nimeth et al., 2004; Pfister et al., 2007; Pfister et al., 2008; De Mulder et al., 2010). Another close relative of *M. lignano*, *Macrostomum hystrix* can also be cultured in the laboratory, and has been used recently as a model for reproduction strategy studies (Ramm et al., 2012). *I. pulchra* with its stem cell system being recently described (De Mulder et al., 2009), is also developing into a useful model organism (Moreno et al., 2010; Chiodin et al., 2013; Perea-Atienza et al., 2013).

Here we report the initial draft genomes of *M. lignano*, *M. hystrix* and *I. pulchra*. The project is ongoing and the results are not final but information generated so far provides insights on how to further proceed with the assemblies of these genomes. We also discuss the rational for choosing these genomes for sequencing and how comparisons of these genomes could improve our understanding of the evolution of bilaterians.

RESULTS

1. *M. lignano* genome assembly

The estimated size of the genome of *M. lignano* is 557 – 655 MB (Ryan Gregory, personal communication). For sequencing we selected isofemale line DV1, which was generated by more than 35 generations of sibling breeding. Initially we generated several shotgun libraries and mate-pair libraries with 3kb, 8kb and 20kb insert size for 454 platform and sequenced these libraries in several runs, producing a total of 6x genome coverage of sequencing data (Table 1). Next to 454 data, we generated shotgun libraries with Illumina platform to the total of 35x coverage (Table 1). These data were combined and a hybrid assembly ML100925 generated using Celera WGS assembler (Miller et al., 2008). This assembly has 116 Megabases in 17,663 scaffolds, with N50 scaffold size of 14,5 kb and the longest scaffold of 18 Mb (Table 2). The scaffolds are comprised of 55,379 contigs with N50 contig size of 2.8 kb. Furthermore, slightly more than 1 million degenerated contigs were produced, comprising 334 Megabases (Table 2). Degenerate contigs in Celera assembler terminology are unitigs that cannot be combined into any scaffold e.g. due to unresolved repeats. In total, the size of ML100925 assembly is 450 Mb, which is close to but somewhat smaller than the estimated genome size, suggesting that some of the genome regions might have been collapsed/overassembled. The ML100925 assembly is the first *M. lignano* draft genome assembly that was made publicly available and can be accessed at http://www.macgenome.com. Despite the fact that the assembly is very fragmented, it proved to be already useful for multiple purposes, such as identification of miRNA genes (see Chapter 6) and cloning of promoter regions for the development of transgenesis methods in *M. lignano*. To visualize transcript structures from *de novo* transcriptome assembly (see Chapter 3) as well as RNA-seq data, we set up a custom version of UCSC genome browser (Kent et al., 2002) for *M. lignano* (Figure 2), which is accessible at http://gb.macgenome.org.

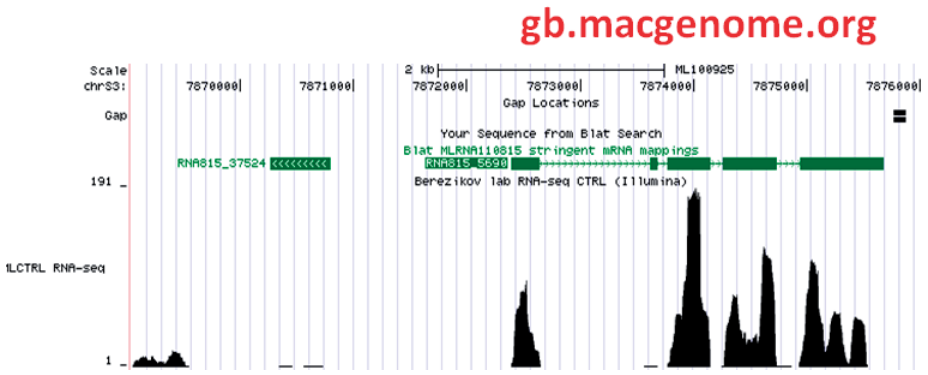


Figure 2. *M. lignano* genome assembly ML100925 is accessible via customized UCSC Genome Browser platform. A screenshot of a transcript and RNA-seq data.

Table 1. Statistics for the genomic libraries used in genome assemblies.

Macrostomum lignano			
library	filtered reads	length	comment
SHG1	1,251,825	332,023,742	454 shotgun
SHG2	1,961,618	633,169,545	454 shotgun
SHG3	2,215,605	728,249,532	454 shotgun
SHG4	1,080,486	350,450,106	454 shotgun
SHG5	1,227,790	424,947,781	454 shotgun
3KB_1	219,785	43,607,257	454 3Kb
3KB_2	234,763	46,858,248	454 3Kb
3KB_3	1,250,995	259,012,232	454 3Kb
8KB_1	1,700,452	299,532,504	454 8Kb
8KB_2	266,899	34,879,000	454 8Kb
20KB_1	349,619	67,963,897	454 20Kb
20KB_2	332,274	62,235,772	454 20Kb
20KB_3	916,967	187,484,600	454 20Kb
20KB_4	908,502	185,998,230	454 20Kb
Total	13,917,580	3,656,412,446	
Raw coverage at 600 Mb		**6x**	
	filtered paired reads	length	
PE250_1	25,319,838	1,924,307,688	Illumina 2x76, 250 nt insert
PE250_2	15,511,970	1,178,909,720	Illumina 2x76, 250 nt insert
PE250_3	22,818,642	1,734,216,792	Illumina 2x76, 250 nt insert
PE250_4	26,722,080	2,030,878,080	Illumina 2x76, 250 nt insert
Total	90,372,530	6,868,312,280	Illumina 2x76, 250 nt insert
Raw coverage at 600 Mb		**11x**	

Table 1. Statistics for the genomic libraries used in genome assemblies. (*Continued*)

Isodiametra pulchra			
library	filtered paired reads	length	comment
IP180b1	47,256,818	4,725,681,800	Illumina 2x100, 180 nt insert
IP180b2	52,796,170	5,279,617,000	Illumina 2x100, 180 nt insert
IP180b3	70,707,602	7,070,760,200	Illumina 2x100, 180 nt insert
IP400b1	29,418,802	2,941,880,200	Illumina 2x100, 400 nt insert
IP400b2	44,049,682	4,404,968,200	Illumina 2x100, 400 nt insert
IP400b3	48,212,296	4,821,229,600	Illumina 2x100, 400 nt insert
Total	292,441,370	29,244,137,000	
Raw coverage at 800 Mb		**37x**	
Macrostomum hystrix			
library	filtered paired reads	length	comment
Hys180	430,638,722	43,063,872,200	Illumina 2x100, 180 nt insert
Hys300	581,444,826	58,144,482,600	Illumina 2x100, 300 nt insert
Hys1_4kb	21,223,554	2,122,355,400	Illumina 2x100, 4 kb insert
Hys1_5kb	20,660,460	2,066,046,000	Illumina 2x100, 5 kb insert
Hys2_4kb	37,241,602	3,724,160,200	Illumina 2x100, 4 kb insert
Hys2_6kb	42,120,522	4,212,052,200	Illumina 2x100, 6 kb insert
Hys2_7kb	12,703,210	1,270,321,000	Illumina 2x100, 7 kb insert
Total reads	1,146,032,896	114,603,289,600	
Raw coverage at 600 Mb		**191x**	

After generating the initial *M. lignano* genome assembly we subsequently generated more Illumina paired-end and mate-pair data for *M. lignano* genome (data not shown) and various algorithms and genomes assemblies were used to assemble these data. However, so far this did not improve the original assembly significantly, indicating that there might be some intrinsic complications in assembling *M. lignano* genome. Indeed, the analysis of kmer spectrum in the raw shotgun sequencing data shows an unusual distribution of kmer frequencies (Figure 3A). While "regular" moderately repetitive and homozygous genomes are expected to show a single peak as for example in Figure 3D (human chromosome 14), in case of *M. lignano* there are two_additional peaks, suggesting potential problem with genome duplication and/or heterozygosity.

2. *I. pulchra* genome assembly

To initiate comparative genomics studies between flatworms and acoels, which also have high regeneration capacity, we performed initial sequencing of *Isodiametra pulchra* genome using purely Illumina sequencing platform. The estimated genome size of *I. pulchra* is about 800 Mb (Ryan Gregory, personal communication), and we made paired-end libraries with

insert sizes of 180 nt and 400 nt in order to utilize ALLPATH-LG assembly pipeline (), which has very specific requirements for the input data. In total, 37 genome coverage of raw data were produced (Table 1), however, ALLPATH-LG assembler failed to generate a useful assembly from it, as well as several other de Bruijn-graph assemblers we tried (data not shown). Instead, we again used Celera WGS assembler and obtained assembly IP120716 (Table 2). This assembly has 648 Mb of sequence but the majority of it, 553 Mb lays in generate contigs, while only 95 Mb were scaffolded, with N50 scaffold size of 9 kb and N50 contig size of 6.7 kb (Table 2). Since no long-range mate pair data were used in this assembly, the small scaffold size is expected. Similar to the *M. lignano* assembly, despite its high fragmentation, this assembly was useful for identification of miRNAs in *I. pulchra* (this work is out of scope of this thesis, however, and not presented here).

Unlike in *M. lignano,* kmer distribution analysis does not show multiple peaks in raw *I. pulchra* data (Fig. 3B). However, there is also no single prominent peak, suggesting that this is not a straightforward genome either. We are currently in the process of generating additional long-range mate-pair data for this genome, which hopefully will help with increasing the continuity of the assembly.

3. *M. hystrix* genome assembly

Taking advantage of continuing developments in the high-throughput sequencing field, very recently we also performed sequencing of *M. hystrix* genome using the new Nextera library preparation protocols for making long-range libraries. We made paired end libraries with 180 nt and 300 nt insert sizes, as well as mate pair libraries with inserts of 4 kb, 6 kb and 7 kb. In total, we generated 191x raw base coverage when calculated for a 600 Mb genome size (there is no experimentally estimate genome size available for *M. hystrix*). Using these data ALLPATH-LG generated a 147 Mb assembly with contig size of 5 kb and scaffold size of 7 kb (data not shown). These results were not particularly impressive and hence we again resorted to Celera WGS assembler. Since Celera assembler is not particularly efficient for high coverage Illumina data, we first assembled paired-end reads into so-called pseudo-Sanger (PS) reads using recently developed anytag assembler (Ruan et al., 2013). Anytag performs a local assembly of paired-end reads and fills the gaps between the ends, thus generating longer and nearly error-free sequences with the characteristics very similar to Sanger reads. This approach was demonstrated to improve the quality of genome assemblies substantially but requires paired end-libraries of 200 nt, 300 nt, 400 nt and 600 nt for best performance (Ruan et al., 2013). While we did not have the whole range of these sized, we successfully used anytag to combine our 180 nt and 300 nt libraries into PS reads of length in the range between 230 nt and 350 nt. These PS reads were provide to Celera WGS assembler together with mate-pair data, and *M. hystrix* genome assembly MH140114 was generated (Table 2). The size of this assembly is 307 Mb, of which 126 Mb are in scaffolds, and scaffold N50 size is 34,7 kb – a noticeable improvement compared to our other genome assemblies. Still, the assembly is rather fragmented, and kmer distribution analysis show three distinct peaks (Fig. 3C), suggesting that similar to *M. lignano* assembly of this genome might be confined by genome duplications and heterozygosity.

Table 2. Assembly statistics for three genomes.

		M. lignano ML100925	*I. pulchra* IP120716	*M. hystrix* MH140114
[Scaffolds]	TotalScaffolds	17,663	14,799	15,921
	TotalContigsInScaffolds	55,379	20,481	52,528
	MeanContigsPerScaffold	3.14	1.38	3.30
	MinContigsPerScaffold	1	1	1
	MaxContigsPerScaffold	9,048	22	2,894
	TotalBasesInScaffolds	**116,434,585**	**94,833,309**	**126,196,506**
	MeanBasesInScaffolds	6,592	6,408	7,926
	MinBasesInScaffolds	40	82	48
	MaxBasesInScaffolds	18,303,428	62,899	4,572,563
	N25ScaffoldBases	157,233	15,473	1,082,000
	N50ScaffoldBases	**14,521**	**9,048**	**34,715**
	N75ScaffoldBases	5,036	5,123	9,790
[Contigs]	TotalContigsInScaffolds	55,379	20,481	52,528
	MeanContigLength	2,103	4,630	2,402
	MinContigLength	31	30	40
	MaxContigLength	21,949	49,181	3,150,872
	N25ContigBases	3,951	11,095	25,941
	N50ContigBases	**2,794**	**6,669**	**3,518**
	N75ContigBases	1,930	3,953	1,726
[DegenContigs]	TotalDegenContigs	1,069,026	4,059,567	616,785
	DegenContigLength	**333,639,415**	**553,150,175**	**180,687,920**
	MeanDegenContigLength	312	136	293
	MinDegenContig	40	30	40
	MaxDegenContig	2,605	1,995	1,366,731
[Contigs + DegenContigs]	**Assembly size**	**450,074,000**	**647,983,484**	**306,884,426**
[gcContent]	Content	44.57%	44.77%	49.81%

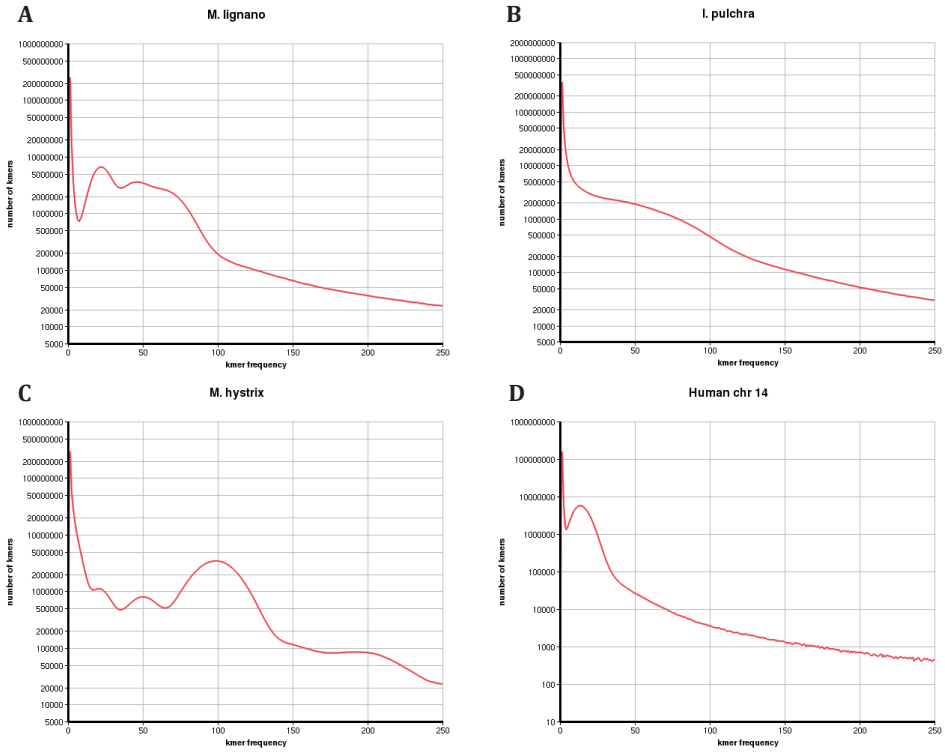


Figure 3. Distribution of 21-nt kmer frequencies in raw sequencing data. **(A)** *M. lignano* has two additional nearly merged peaks. **(B)** *I. pulchra* does not have a prominent peak. **(C)** *M. hystrix* has three regularly spaced peaks at ~25,50 and 100 and a slight elevation at 200. **(D)** An example of kmer distribution in human chromosome 14 data from Genome Assembly Gold-Standard project (Salzberg et al., 2012), which has a single prominent peak.

DISCUSSION AND FUTURE DIRECTIONS

1. Assemblies quality

Here we present initial genome assemblies of two flatworms and an acoel. While advances in next-generation sequencing technologies now make it feasible and affordable generating hundreds of x of genome coverage sequencing data, assembly of the genomes towards the high-quality level often remains problematic and may significantly depend on the structure of a particular genome and such factors the levels of polymorphisms and repeats. In case of *M. lignano* and *M. hystrix* genomes analysis of kmer frequencies in raw sequencing data shows unusual distributions (Fig 3), suggesting that there might be partial genome duplications and significant levels of polymorphisms in these genomes. Indeed, very recently we initiated cytogenetic study in *M. lignano* and discovered that DV1 line used for genome sequencing contains an additional pair of chromosomes. While wild-type *M. lignano* has 4 pairs of chromosomes (Egger and Ishida, 2005), in DV1 there are 5 pairs, whit the largest chromosome apparently duplicated (Zadesenets

et al, in preparation). This very recent duplication might explain part of the observed kmer profile but it also suggest that *Macrostomum* is very flexible with chromosome dynamics and thus more and older duplicated regions are likely to be present in its genome. The similar situation is likely to be with *M. hystrix* genome. One potential way to overcome these difficulties is by using long PacBio reads to resolve repeats and shorter duplications. An effort is currently ongoing in this direction and in the future PacBio data will be used to improve *M. lignano* genome assembly. Furthermore, the pseudo-Sanger read approach (Ruan et al., 2013), combined with long-range mate-pair libraries seems promising based on our results with *M. hystrix* genome assembly and should be extended to the genomes of *M. lignano and I. pulchra* as well.

At the second stage of this project we plan to use all the produced data from these three species for large-scale genomic comparative analysis. *M. lignano* is one of the most basal bilaterians (Ehlers, 1985; Riutort et al., 2012). Evolutionary position of *I. pulchra* is still under debate with two most popular versions putting it either at the very base of bilaterian (Egger et al., 2009) or among the oldest deuterostomes (Philippe et al., 2011). In both cases comparing gene complements of these two species would allow reconstructing the gene repertoire of the common bilaterian ancestor. If acoels are the most basal bilaterians, than such analysis would mean the comparison of two oldest bilaterian phyla, and thus provide the most valuable and detailed insight into the gene set of the bilaterian ancestor. And if *I. pulchra* is more derived and represents the most basal deuterostomes, it would help to identify the genes possibly involved in emergence of deuterostomal way of embryonic development. In any case genomes of *M. lignano* and *I. pulchra* would add greatly to our understanding of the evolution of bilaterians.

Even more importantly, *M. lignano* and *I. pulchra* represent the only two phyla within animals with bilateral symmetry that possess adult pluripotent stem cell populations (Sato et al., 2006; De Mulder et al., 2009; Bely and Sikes, 2010; Baguñà, 2012). Genes shared between the two species, but not present in higher metazoans (for example, human) would potentially be the ones responsible for the establishment and maintenance of this unique stem cell system and impressive regeneration capacity. It is of extreme interest to check whether those genes can also be found in the recently sequenced genomes of cnidarians (*N. vectensis* or *Hydra*) or sponges (*A. queenslandica*) in order to better understand the evolution of pluripotency.

Comparison of the gene repertoires of *M. lignano* and *N. vectensis* can help to reconstruct the genome of the eumetazoan ancestor, common for bilaterians and lower animals with radial symmetry. Such an attempt was already made, based on the comparative analysis of the gene sets of sea anemone and human (Putnam et al., 2007). *Macrostomum*, however, is way more basal organism than *Homo sapiens*; thus, its genome structure should be much closer to the ancestral version. We assume that human has gained or lost much more genes in the course of its evolution than one of the most basal flatworms, so the reconstruction of the ancestral gene set based on the comparison of *N. vectensis* and *M. lignano* would be more complete. We also plan to address the conservation between this reconstructed ancestral genome and the genomes of highly derived eumetazoans such as nematodes or fruit flies.

The next step would be to compare two hypothetical genomes – of the common eumetazoan and common bilaterian ancestor. Such comparison would provide us with the list of candidate genes potentially involved in the emergence of bilateral symmetry.

M. lignano and *M. hystrix* are seemingly highly similar species. There are, however quite significant differences – for example, in mating behavior or the ability of *M. hystrix* to self-fertilize, which lack is in *M. lignano* (Ramm et al., 2012). Comparison of the gene sets of these species would help to estimate how close are they evolutionary. Besides, having detailed information about both closely related genomes would help to develop *M. hystrix* into the genetically complementary model to *M. lignano* – great tool to address evolutionary questions.

CONCLUSION

We have initiated genome sequencing of the three species – *M. lignano, I pulchra* and *M. hystrix* – and used different Next Generation Sequencing platforms for this project. We presented draft assemblies of all three genomes. *M. lignano* genome assembly version ML100925 is publicly available for downloading or blast search at http://www.macgenome.org. The other two genomes would be available shortly. The quality of the current assemblies is satisfactory and allows addressing some of the questions, such as miRNA annotation, identification of promoter regions and establishment of gene models. The genome assemblies still required significant improvements and the work in this direction will continue.

Better assemblies would allow us to perform large-scale comparative analysis. These comparisons would, in turn, shed light on many secrets, such as evolutionary position of acoels, emergence of bilateral symmetry and evolution of the mechanisms controlling pluripotent adult stem cells. Additionally we plan to use gene repertoires of *M. lignano*, *I. pulchra* and *N. vectensis* to reconstruct the gene set of common bilaterian ancestor and improve genetic reconstruction of common eumetazoan ancestor, reported previously.

MATERIALS AND METHODS

1. Animal culture

M. lignano, M. hystrix and *I. pulchra* animals were cultured in Petri dishes in f/2 medium, nutrient-enriched artificial sea water (Andersen et al., 2005), and were fed with unicellular diatom *Nitzschia curvilineata* (Heterokontophyta, Bacillariophyceae), which covers the bottom of the petri dishes (Ladurner et al., 2005). Medium for I. pulchra was additionally supplied with kanamycin (50 µg/ml). All animals were kept in the incubators under following conditions: 20°C temperature and 60% humidity. *Macrostomum* species were kept at 14/10 hours day/night cycle (Rieger et al., 1988). *I. pulchra* cultures were maintained in the absence of light. In case of all three species mixed populations were used for DNA extraction.

2. Genomic DNA extraction

During a week prior to DNA isolation animals were kept on antibiotic-containing medium. Medium was changed every day with 50 µg/ml streptomycin or ampicillin added (antibiotics were also switched every day). Worms were starved 24 hours prior to extraction, and then rinsed in the fresh medium. Genomic DNA was extracted using the USB PrepEase Genomic DNA Isolation kit (USB-Affymetrix, Cat. No. 78855) according to manufacturer's instructions. For the lysis step worms were kept in the supplied lysis buffer (with Proteinase K added) at 55°C for

30-40 minutes and mixed by inverting the tube every 5 minutes. DNA was ethanol-precipitated once following the extraction and resuspended in TE buffer (for making 454 libraries Qiagen EB buffer was used instead). Concentration of DNA samples was measured with Qubit dsDNA BR assay kit (Life Technologies, Cat. No. Q32850).

3. Preparation and sequencing of 454 genomic DNA libraries

454 shotgun DNA libraries were made with GS FLX Titanium General Library Preparation Kit (Roche, Cat. No. 05233747001, and for paired end libraries the set of GS FLX Titanium Library Paired End Adaptors (Roche, Cat. No. 05463343001) was additionally used. All the libraries were made following the manufacturer's protocol and sequenced on454 FLX and FLX Titanium systems.

4. Preparation and sequencing of Illumina genomic DNA libraries

The standard Illumina kits for preparation of paired-end genomic libraries were used following the manufacturer's protocol and the libraries were sequenced on Illumina Genome Analyzer IIx and HiSeq2500 systems. Long-range mate-pair libraries were prepared with Nextera Mate Pair Sample Preparation Kit (Illumina, Cat No. FC-132-1001) according to manufacturer's protocol.

5. Genome assembly

Genome assembly strategies are described in details in the results section.

REFERENCES

Andersen, R.A., Berges, R.A., Harrison, P.J., and Watanabe, M.M. (2005). Recipes for freshwater and seawater media; enriched natural seawater media. In: Anderson, R.A. (Ed.), Algal culturing techniques, Elsevier academic press, Burlington, 507.

Baguñà, J. (2012). The planarian neoblast: the rambling history of its origin and some current black boxes. Int J Dev Biol *56*, 19-37.

Baguñà, J., and Riutort, M. (2004). Molecular phylogeny of the Platyhelminthes. Canadian journal of zoology *82*, 168-193.

Bely, A.E., and Sikes, J.M. (2010). Acoel and platyhelminth models for stem-cell research. J Biol *9*, 14.

Benfey, P.N. (1999). Stem cells: A tale of two kingdoms. Curr Biol *9*, R171-72.

Bode, A., Salvenmoser, W., Nimeth, K., Mahlknecht, M., Adamski, Z., Rieger, R.M., Peter, R., and Ladurner, P. (2006). Immunogold-labeled S-phase neoblasts, total neoblast number, their distribution, and evidence for arrested neoblasts in Macrostomum lignano (Platyhelminthes, Rhabditophora). Cell Tissue Res *325*, 577-587.

Bode, H.R. (1996). The interstitial cell lineage of hydra: a stem cell system that arose early in evolution. J Cell Sci *109 (Pt 6)*, 1155-164.

Bosch, T.C. (2008). Stem cells (Springer).

Bosch, T.C., Anton-Erxleben, F., Hemmrich, G., and Khalturin, K. (2010). The Hydra polyp: nothing but an active stem cell community. Dev Growth Differ *52*, 15-25.

Chapman, J.A., Kirkness, E.F., Simakov, O., Hampson, S.E., Mitros, T., Weinmaier, T., Rattei, T., Balasubramanian, P.G., Borman, J., and Busam, D. (2010). The dynamic genome of Hydra. Nature *464*, 592-96.

Chiodin, M., Børve, A., Berezikov, E., Ladurner, P., Martinez, P., and Hejnol, A. (2013). Mesodermal gene expression in the acoel Isodiametra pulchra indicates a low number of mesodermal cell types and the endomesodermal origin of the gonads. PLoS One *8*, e55499.

Darling, J.A., Reitzel, A.R., Burton, P.M., Mazza, M.E., Ryan, J.F., Sullivan, J.C., and Finnerty, J.R. (2005). Rising starlet: the starlet sea anemone, Nematostella vectensis. Bioessays *27*, 211-221.

Demircan, T., and Berezikov, E. (2013). The Hippo pathway regulates stem cells during homeostasis and regeneration of the flatworm Macrostomum lignano. Stem Cells Dev *22*, 2174-185.

Edgecombe, G.D., Giribet, G., Dunn, C.W., Hejnol, A., Kristensen, R.M., Neves, R.C., Rouse, G.W., Worsaae,

K., and Sørensen, M.V. (2011). Higher-level metazoan relationships: recent progress and remaining questions. Organisms Diversity & Evolution *11*, 151-172.

Egger, B., and Ishida, S. (2005). Chromosome fission or duplication in Macrostomum lignano (Macrostomorpha, Plathelminthes)--remarks on chromosome numbers in 'archoophoran turbellarians'. Journal of Zoological Systematics and Evolutionary Research *43*, 127-132

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

Egger, B., Steinke, D., Tarui, H., De Mulder, K., Arendt, D., Borgonie, G., Funayama, N., Gschwentner, R., Hartenstein, V., et al. (2009). To be or not to be a flatworm: the acoel controversy. PLoS One *4*, e5502.

Ehlers, U. (1985). phylogenetische System der Plathelminthes (G. Fischer).

Ehlers, U., and Sopott-Ehlers, B. (1995). Plathelminthes or platyhelminthes? Hydrobiologia *305*, 1-2.

Funayama, N. (2010). The stem cell system in demosponges: insights into the origin of somatic stem cells. Dev Growth Differ *52*, 1-14.

Funayama, N. (2013). The stem cell system in demosponges: suggested involvement of two types of cells: archeocytes (active stem cells) and choanocytes (food-entrapping flagellated cells). Dev Genes Evol *223*, 23-38.

Galliot, B. (2012). Hydra, a fruitful model system for 270 years. Int J Dev Biol *56*, 411-423.

Gschwentner, R., Ladurner, P., Nimeth, K., and Rieger, R. (2001). Stem cells in a basal bilaterian. Cell Tissue Res *304*, 401-08.

Hemmrich, G., and Bosch, T.C. (2008). Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. Bioessays *30*, 1010-18.

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D.B., and Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution *67*, 3233-242.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006

Ladurner, P., Pfister, D., Seifarth, C., Schärer, L., Mahlknecht, M., Salvenmoser, W., Gerth, R., Marx, F., and Rieger, R. (2005). Production and characterisation of cell- and tissue-specific monoclonal antibodies for the flatworm Macrostomum sp. Histochem Cell Biol *123*, 89-104.

Ladurner, P., Rieger, R., and Baguñà, J. (2000). Spatial distribution and differentiation potential of stem cells in hatchlings and adults in the marine platyhelminth macrostomum sp.: a bromodeoxyuridine analysis. Dev Biol *226*, 231-241.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Maxmen, A. (2011). Evolution: A can of worms. Nature *470*, 161-62.

Miller, J.R., Delcher, A.L., Koren, S., Venter, E., Walenz, B.P., Brownley, A., Johnson, J., Li, K., Mobarry, C., and Sutton, G. (2008). Aggressive assembly of pyrosequencing reads with mates. Bioinformatics *24*, 2818-824.

Moreno, E., De Mulder, K., Salvenmoser, W., Ladurner, P., and Martínez, P. (2010). Inferring the ancestral function of the posterior Hox gene within the bilateria: controlling the maintenance of reproductive structures, the musculature and the nervous system in the acoel flatworm Isodiametra pulchra. Evol Dev *12*, 258-266.

Mouton, S., Willems, M., Braeckman, B.P., Egger, B., Ladurner, P., Schärer, L., and Borgonie, G. (2009). The free-living flatworm Macrostomum lignano: a new model organism for ageing research. Exp Gerontol *44*, 243-49.

De Mulder, K. (2009). Molecular characterization of the exceptional stem cell system of flatworms. PhD thesis, Ghent University.

De Mulder, K., Kuales, G., Pfister, D., Egger, B., Seppi, T., Eichberger, P., Borgonie, G., and Ladurner, P. (2010). Potential of Macrostomum lignano to recover from gamma-ray irradiation. Cell Tissue Res *339*, 527-542.

De Mulder, K., Kuales, G., Pfister, D., Willems, M., Egger, B., Salvenmoser, W., Thaler, M., Gorny, A.K., Hrouda, M., et al. (2009). Characterization of the stem cell system of the acoel Isodiametra pulchra. BMC Dev Biol *9*, 69.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., et al. (2009). Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. Dev Biol *334*, 198-212.

Müller, W.E. (2006). The stem cell concept in sponges (Porifera): Metazoan traits. Semin Cell Dev Biol *17*, 481-491.

Nimeth, K.T., Mahlknecht, M., Mezzanato, A., Peter, R., Rieger, R., and Ladurner, P. (2004). Stem cell dynamics during growth, feeding, and starvation in the basal flatworm Macrostomum sp. (Platyhelminthes). Dev Dyn *230*, 91-99.

Perea-Atienza, E., Botta, M., Salvenmoser, W., Gschwentner, R., Egger, B., Kristof, A., Martinez, P., and Achatz, J.G. (2013). Posterior regeneration in Isodiametra pulchra (Acoela, Acoelomorpha). Front Zool *10*, 64.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. (2007). The exceptional stem cell system of Macrostomum lignano: screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. Front Zool *4*, 9.

Philippe, H., Brinkmann, H., Copley, R.R., Moroz, L.L., Nakano, H., Poustka, A.J., Wallberg, A., Peterson, K.J., and Telford, M.J. (2011). Acoelomorph flatworms are deuterostomes related to Xenoturbella. Nature *470*, 255-58.

Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., et al. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science *317*, 86-94.

Ramm, S.A., Vizoso, D.B., and Schärer, L. (2012). Occurrence, costs and heritability of delayed selfing in a free-living flatworm. J Evol Biol *25*, 2559-568.

Rieger, R.M., Legniti, A., Ladurner, P., Reiter, D., Asch, E., Salvenmoser, W., Schurmann, W., and Peter, R. (1999). Ultrastructure of neoblasts in microturbellaria: significance for understanding stem cells in free-living Platyhelminthes. Invertebrate reproduction & development *35*, 127-140.

Riutort, M., Álvarez-Presas, M., Lázaro, E., Solà, E., and Paps, J. (2012). Evolutionary history of the Tricladida and the Platyhelminthes: an up-to-date phylogenetic and systematic account. Int. J. Dev. Biol *56*, 5-17.

Ruan, J., Jiang, L., Chong, Z., Gong, Q., Li, H., Li, C., Tao, Y., Zheng, C., Zhai, W., et al. (2013). Pseudo-Sanger sequencing: massively parallel production of long and near error-free reads using NGS technology. BMC Genomics *14*, 711.

Salzberg, S.L., Phillippy, A.M., Zimin, A., Puiu, D., Magoc, T., Koren, S., Treangen, T.J., Schatz, M.C., Delcher, A.L., et al. (2012). GAGE: A critical evaluation of genome assemblies and assembly algorithms. Genome Res *22*, 557-567.

Sato, K., Shibata, N., Orii, H., Amikura, R., Sakurai, T., Agata, K., Kobayashi, S., and Watanabe, K. (2006). Identification and origin of the germline stem cells as revealed by the expression of nanos-related gene in planarians. Dev Growth Differ *48*, 615-628.

Simanov, D., Mellaart-Straver, I., Sormacheva, I., and Berezikov, E. (2012). The Flatworm Macrostomum lignano Is a Powerful Model Organism for Ion Channel and Stem Cell Research. Stem Cells Int *2012*, 167265. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M.E., Mitros, T., Richards, G.S., Conaco, C., Dacre, M., et al. (2010). The Amphimedon queenslandica genome and the evolution of animal complexity. Nature *466*, 720-26.

Verdoodt, F., Bert, W., Couvreur, M., De Mulder, K., and Willems, M. (2012). Proliferative response of the stem cell system during regeneration of the rostrum in Macrostomum lignano (Platyhelminthes). Cell Tissue Res *347*, 397-406.

Watanabe, H., Hoang, V.T., Mättner, R., and Holstein, T.W. (2009). Immortality and the base of multicellular life: Lessons from cnidarian stem cells. Semin Cell Dev Biol *20*, 1114-125.

Zhan, M., Miura, T., Xu, X., and Rao, M.S. (2005). Conservation and variation of gene regulation in embryonic stem cells assessed by comparative genomics. Cell Biochem Biophys *43*, 379-405.

CHAPTER 3

DE NOVO TRANSCRIPTOME ASSEMBLY OF *MACROSTOMUM LIGNANO* AND IDENTIFICATION OF NEOBLAST-SPECIFIC GENES

Daniil Simanov[1,*], Stijn Mouton[1,2,*], Wibowo Arindrarto[1], Kay van Nies[1], Katrien de Mulder[1], Magda Grudniewska[2], Turan Demircan[1], Gerard van der Krogt[1], Birgit Lengerer[3], Peter Ladurner[3] and Eugene Berezikov[1,2].

*These authors contributed equally to this work.

[1]Hubrecht Institute, KNAW and University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands
[2]European Research Institute for the Biology of Ageing, University of Groningen, University Medical Centre Groningen, 9713AV Groningen, The Netherlands
[3]Ultrastructural Research and Evolutionary Biology, Institute of Zoology, University of Innsbruck, Technikerstrasse 25, 6020 Innsbruck, Austria

Manuscript in preparation.

ABSTRACT

Marine flatworm *Macrostomum lignano* is an attractive model for stem cell studies. It possesses extraordinary regeneration capabilities mediated by a population of adult somatic stem cells known as neoblasts. Elucidation of *M. lignano* transcriptome and transcriptional landscape of its stem cells will increase our understanding of the mechanisms that regulate stem cell functioning, tissue homeostasis and regeneration. We generated *de novo* transcriptome assembly of *M. lignano* and used it as a basis for characterization of transcriptional dynamics in *M. lignano* over a time course after irradiation and at different stages of post-embryonic development. We applied different selection strategies and cutoffs to the generated RNA-seq data and compiled three alternative lists of candidate neoblast-enriched genes. We validated their expression in dividing cells by *in situ* hybridization. Lastly, we addressed the function of selected candidate genes by RNA interference and identified several neoblast-related phenotypes. Our candidate lists contain almost all of the previously described neoblast-specific genes and are enriched for nuclear components, chromatin remodeling and cell division factors. The most interesting novel neoblast-specific gene identified is evolutionary conserved helicase *DDX39*. Knockdown of this gene by RNA interference gene leads to fast block of neoblast proliferation and subsequent severe failure of tissue homeostasis and regeneration, while overexpression or knockdown of human *DDX39* in cell cultures drastically changes proliferation rate *in vitro*. Our data show that *M. lignano* is a convenient model for studying different aspects of stem cell biology, and its *de novo* assembled transcriptome can be successfully used as a tool for identifying genes involved in different aspects of stem cell functioning.

INTRODUCTION

Pluripotent stem cells represent potentially unlimited *in vitro* source of every cell type, which makes them particularly interesting not only for fundamental studies, but also for regenerative medicine. Vertebrate adults do not have pluripotent somatic stem cells, but pluripotent cells can be derived from early embryos (Weissman, 2000) or via reprogramming of somatic cells (Takahashi and Yamanaka, 2006). Understanding the basis of cellular pluripotency and controlled transition to progenitor differentiation, as well as the migration of stem cells within an organism is of the outmost importance as stem cell therapies are being quickly developed (Lin et al., 2013; Okano et al., 2013). These therapies have been in use for nearly half a century, quickly advancing to clinical trials regardless of our poor understanding of the mechanisms controlling different aspects of stem cell functioning. Thus, better understanding of the processes that specifically occur in adult stem cells would have a huge impact on regenerative medicine, improving the efficacy of therapies and possibly providing novel approaches to combat degenerative diseases and repair injuries.

Today a major limitation in the study of somatic stem cells lies in the difficulty of accessing these cells *in vivo*, especially in vertebrates. Thus, simple model organisms with experimentally accessible and actively functioning stem cells should be very useful in identifying and functionally testing the mechanisms regulating stem cell activities. To date only a limited number of models that can fill this experimental gap are available for the scientists, and flatworms are the most well-studied and well-established among them. Flatworms possess an extraordinary stem cell system that can give rise to virtually all cell types during development and regeneration. These stem cells – called neoblasts – are the only somatic cells that divide in the adult and are responsible for the enormous regeneration capacity and tissue plasticity (Agata et al., 2006; Rink, 2013). Since the end of the 19th century, when regeneration of the species *Planaria maculata* and *Planaria lugubris* were first described by Thomas Hunt Morgan (Morgan, 1898; Morgan, 1902), planarian flatworms have been used by a great number of groups, becoming a traditional model for stem cell and regeneration research. Today two planarian species most commonly used in laboratories are *Schmidtea mediterranea* and *Dugesia japonica* (Sánchez Alvarado, 2004; Agata et al., 2006; Gentile et al., 2011; Baguñà, 2012).

During the last 15 years development of new techniques and experimental methods has sparked planarian research. The genome database of *S. mediterranea* is publicly available (Robb et al., 2008) and a genome project of *D. japonica* has been initiated. The current planarian toolkit further includes robust in situ hybridization (ISH) (Umesono et al., 1997; Pearson et al., 2009), BrdU–labeling (Newmark and Sánchez Alvarado, 2000) techniques and FACS fractionation of cell populations (Hayashi et al., 2006; Hayashi and Agata, 2012). Gene-specific knockdowns by RNA interference (RNAi) (Sánchez Alvarado and Newmark, 1999; Newmark et al., 2003; Reddien et al., 2005) made it possible to identify signal transduction pathways and specific genes that regulate different aspects of regeneration (Gurley et al., 2008; Petersen and Reddien, 2008; Adell et al., 2009; Rink et al., 2009; Felix and Aboobaker, 2010; Oviedo et al., 2008) and stem cell functioning (Oviedo et al., 2008; Guo et al., 2006; Fernandéz-Taboada et al., 2010; Scimone et al., 2010). Exposure of animals to lethal doses of γ-irradiation together with improved transplantation techniques allowed Wagner and colleagues to demonstrate the real pluripotency (and likely totipotency) of planarian neoblasts (Wagner et al., 2011).

Combination of all the methods mentioned above with next generation sequencing techniques made planarians an outstanding tool for gene discovery (Friedländer et al., 2009; Solana et al., 2012; Onal et al., 2012). In recent years a number of independent transcriptome assemblies for *S. mediterranea* have been completed (Blythe et al., 2010; Adamidi et al., 2011; Sandmann et al., 2011; Nishimura et al., 2012), and different approaches have been used to identify genes, both well-known and novel, specifically involved in proliferation and regulation of planarian stem cells (Friedländer et al., 2009; Solana et al., 2012; Onal et al., 2012; Blythe et al., 2010; Rossi et al., 2007; Eisenhoffer et al., 2008; Wagner et al., 2012). In the long list of genes now known to be expressed in neoblasts, one can pinpoint a few groups clearly standing out: components of cell division machinery, proteins involved in posttranscriptional regulation of gene expression, and chromatin modifiers (Solana et al., 2012; Onal et al., 2012; Rossi et al., 2007; Eisenhoffer et al., 2008; Wagner et al., 2012; Shibata et al., 2012; Labbé et al., 2012; Rouhana et al., 2010).

Planarian flatworms as model systems contributed a lot to our understanding of stem cell functioning, yet currently there are technical limits to what we could learn from these animals, and many important questions in the field remain unanswered. Experiments of Wagner and colleagues demonstrated that at least 5% of planarian neoblasts are indeed pluripotent, but also that some are clearly not (Wagner et al., 2011). This unknown degree of population heterogeneity amongst neoblasts remains one of the biggest problems in planarian stem cell research, both conceptually and experimentally (Rink, 2013). We cannot yet define these stem cell subpopulations and still know little about the mechanisms that regulate them during growth and regeneration, maintaining and re-establishing the dynamic steady state between neoblasts and differentiated cells. Almost nothing is known about embryonic origins of adult somatic stem cells and regulation of neoblast divisions at the single cell level (Rink, 2013; Baguñà, 2012). More information, detailed transcription profiling of stem cell subpopulations and, most importantly, new tools such as transgenics are required for further investigation of all these problems.

Another intriguing question concerning neoblasts is evolution of pluripotency. Two recent studies compared neoblasts to vertebrate stem cells, reporting a high degree of overlap between pluripotency associated set of genes expressed in vertebrate stem cells and neoblast-specific genes (Onal et al., 2012; Labbé et al., 2012). Yet comparison of pluripotency across shorter evolutionary distances still needs to be done, and studying neoblasts or their analogues in other invertebrate systems provides an opportunity to approach the evolution of animal stem cell systems.

Today another flatworm model, *Macrostomum lignano* can be successfully used to expand our knowledge about neoblasts and complement the research done on planarian worms (Ladurner et al., 2005; Egger et al., 2006). *M. lignano* is marine free-living flatworm, about 1,5 mm long, that consists of roughly 25000 cells. Around 6-7 % of these cells are adult somatic stem cells, which facilitate regeneration capacity of the animal. All the neoblasts are located mesodermally, in a bilateral pattern and never anterior from the eye level. Proliferating cells can be also found in gonads, and somatic stem cells can give origin to germ line (Fig. 1B) (Egger et al., 2006). Transparency of the worm makes it possible to study virtually any processes in a living

specimen under the microscope (Fig. 1A). Current toolkit available for work with *M. lignano* includes BrdU labeling by soaking, robust ISH technique, set of tissue-specific monoclonal antibodies and simple RNAi procedure, also by soaking (Pfister et al., 2007; Pfister et al., 2008; Ladurner et al., 2005). Importantly, *M. lignano* is non-self-fertilizing hermaphrodite and has exclusively sexual reproduction. Well-fed adult animal generate a lot of eggs all year through (one animal lays one egg a day on average), making it amenable for genetic manipulation. Unlimited access to single eggs during the whole year also allows the analysis of the embryonic development (Morris et al., 2004; Morris et al., 2006). The draft genome of *M. lignano* is sequenced and annotated (see http://www.macgenome.org and Chapter 2 of this Thesis), and, most importantly, first transgenic lines are available (Fig. 1C) (Demircan et al, manuscript in preparation). With all these characteristics and well-developed toolbox *M. lignano* is a promising model system for *in vivo* stem cell biology, complementary to planarian flatworms.

Here we present first *de novo* assembly of the transcriptome of *M. lignano* and demonstrate that it can be successfully used for gene expression profiling studies using RNA-seq. We compare expression levels of genes in intact worms with the ones that were irradiated in order to remove proliferating cells. We used three different time points after irradiation to collect mRNA from the worms and study dynamical changes in gene expression levels after all the dividing cells were removed. We also compared transcriptomes from different stages of development. Applying different selection criteria and cutoffs to these sets of RNA-seq data we defined three alternative sets of genes whose expression is potentially enriched in somatic stem cells and their progeny. In this study we only focused on evolutionary conserved genes,

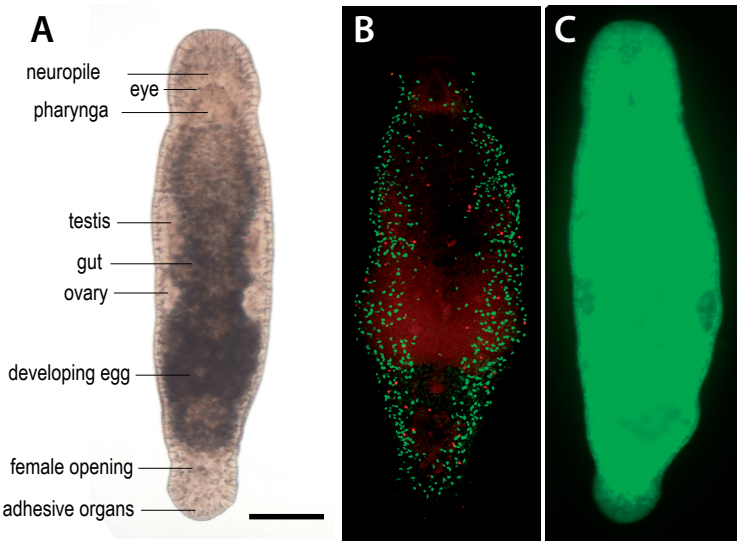


Figure 1. *Macrostomum lignano* as a model organism. **(A)** Bright field image of a living specimen. **(B)** Confocal projection of BrdU and phospho- histone H3 immunostaining after 30 minutes BrdU pulse in an adult worm (green: S-phase cells, red: mitotic cells). **(C)** Hub1 transgenic worm (characterized by ubiquitous GFP expression) under fluorescence microscope. Scalebar 150 μm (all worms).

ruling out those specific for *M. lignano*. We validated expression of our candidate transcripts in dividing cells by ISH and addressed function of selected genes by RNAi. We describe nine quickly developing phenotypes, associated with the loss or malfunction of proliferating cells.

All three of our candidate lists contain for the most part previously described neoblast-associated genes and are enriched for the three groups of genes mentioned above – cell division factors, chromatin modifiers and posttranscriptional regulators. We found a significant number of novel candidates as well, and RNAi on one of them, DEAD box RNA helicase *DDX39*, caused one of the fastest and the most severe phenotypes in our screen.

RESULTS

1. Sequencing and de novo assembly of *M. lignano* transcriptome

During the course of this study we generated two *de novo* transcriptome assemblies of *M. lignano*, which reflects the general developments in the high-throughput sequencing field in the recent years, with different sequencing platforms becoming more accessible and robust with each year, as well as constant improvements in assembly algorithms. We released the first public *M. lignano* transcriptome assembly, called MLRNA110815, in August of 2011, and it was based on combination of data generated by 454, Illumina and SOLiD platforms. More recently, we generated additional substantial amount of strand-specific Illumina RNA-seq data and released an improved transcriptome assembled MLRNA131024. Here we describe how both assemblies were generated and note that the neoblast gene candidate gene screen described in this Chapter is based on the earlier transcriptome assembly version for historical reasons, whereas the analysis of gene expression during adaptation to low salinity, described in Chapter 5, is based on the newer assembly version.

***De novo* transcriptome assembly MLRNA110815.** Initially we generated RNA-seq data from mixed population of animals using 454 sequencing platform, and cDNA for the 454 library preparation was normalized by subtractive hybridization prior to sequencing in order to increase chances of detecting low-abundant transcripts. In total, 1.1 mln reads with an average length of 300 nt were generated on 454 platform (Table 1). Next, we generated RNA-seq data from control animals and animals 7 days after irradiation with a cumulative does of 210 Gray, and the same mRNA material was sequenced using both SOLiD and Illumina platforms (libraries CTRL1, 210G, CTRL1i, 210Gi, Table 1).

These data were used to generate the *de novo* transcriptome assembly in a multi-step protocol (Sup. Fig. 1), using Mira EST assembler as the primary assembly software because it allowed combining data from different sequencing platforms to produce *de novo* transcriptome assembly (Chevreux et al., 2004). Since Mira cannot handle very efficiently large amounts of Illumina data directly, the paired-end Illumina data (2x100 nt libraries CTRL1i and210Gi, Table1) were preprocessed. First, paired ends were merged with custom scripts and cap3 software (Huang and Madan, 1999) to generate 23.4 mln longer merged fragments with an average length of 115 nt (Sup. Fig. 1). These fragments were assembled with the Inchworm assembler from Trinity package (Grabherr et al., 2011) into 791,319 initial transcripts, and next these data were combined with 454 reads and 19,839 available ESTs (Morris et al., 2006) and

provided to Mira assembler. Mira assembled 125,146 contigs from these data (Sup. Fig. 1). Next, data from strand-specific SOLiD libraries (Table 1) were used to verify and correct orientation of the transcripts, and redundancy in the assembly was reduced by removing transcripts that matched to longer transcripts with at least 90% identity and spanning at least 90% of its length. Transcripts shorter than 200 nt were also removed, resulting in 76,437 transcripts in the final MLRNA110815 assembly.

The assembly was further annotated by grouping similar sequences into clusters using cd-hit software (Fu et al, 2012), identifying Pfam domains with HMMER (Finn et al., 2011) and homologues in other species using blastx (Altschul et al., 1997). The 76,437 transcripts in MLRNA110815 assembly cluster into 65,251 transcript groups (Sup. Fig. 1), the N50 size of the assembly is 1,087 nt, and the longest transcript is 27,308 nt (Table 2). 20,482 transcripts have matches in Pfam database (Punta et al., 2012) above inclusion threshold and represent 3,813 Pfam models (Table 2). Furthermore, 24,000 transcripts have human homologs representing 9,151 different human genes (Table 2). Finally, the completeness of a transcriptome assembly can be judged by the coverage of core eukaryotic genes (CEGs) – a set of 248 genes that are present in most eukaryotic species (Parra et al., 2009). In case of MLRNA110815 assembly all but 3 CEGs can be identified (Table 3), and the missing genes belong to more variable ranks (rank 1 and 2), which may be missing in some taxa. At the same time, one of the missing CEGs is from the most conserved rank 4 set (Table 3), suggesting that while the assembly is rather complete, it still may have some missed or misassembled transcripts. Nevertheless, despite the imperfections, we considered MLRNA110815 assembly sufficiently informative to be used for screening candidate neoblast genes as described below.

Table 1. Statistics for RNA-seq libraries used in this work.

Library	raw reads/pairs	used reads/fragments	Platform	Comments
454	1,192,234	1,192,234	454 Titanium	Normalized library from mixed population
CTRL1	18,515,538	5,918,990	SOLiD	Controls for irradiation experiment
210G	19,255,029	5,675,947	SOLiD	Irradiated (7 days after)
CTRL1i	17,827,673	13,683,263	Illumina	Same material as CTRL1
210Gi	13,014,864	9,718,379	Illumina	Same material as 210G
Juv	21,753,873	3,424,939	SOLiD	Juveniles, 4 days after hatching
CTRL2	14,695,410	5,649,482	SOLiD	Controls for Irradiation time course
D1	11,188,631	3,995,757	SOLiD	1 day after irradiation
D3	6,336,404	2,308,571	SOLiD	3 days after irradiation
D7	15,925,394	5,801,432	SOLiD	7 days after irradiation

***De novo* transcriptome assembly MLRNA131024.** Since the RNA-seq field advanced substantially during the course of this work, with improvements in both algorithms for *de novo* transcriptome assembly as well as in methodologies for generating RNA-seq data, we recently attempted to improve *M. lignano de novo* transcriptome assembly taking into account these

recent developments. We generated strand-specific Illumina libraries for investigating salinity adaptation in *M. lignano* (see Chapter 5) and used these data (Table 1) to improve assembly.

Table 2. Statistics for *M. lignano* de novo transcriptome assemblies.

	MLRNA110815	**MLRNA131024**
Contigs	76,437	174,922
Bases	65,656,600	155,512,149
Mean Length	846	889
Min Length	200	200
Max Length	27,308	35,582
N50	1,087	1,343
Pfam annotations	20,482	43,897
Non-redundant Pfam domains	3,813	5,329
Transcripts with human homologs	24,000	52,012
Number of different human homolog genes	9,151	10,268

Our previous experience with MLRNA110815 assembly and experimentation with various assembly algorithms suggested that there is no single best approach for generating *de novo* transcriptome assembly, with different assemblers putting together correct transcripts for different subsets of genes. Similarly, it was reported recently that no single assembly approach can correctly reconstruct a full transcriptome even in a well-studied model, such as human or *Drosophila* (Steijger et al., 2013). We therefore attempted to use different transcriptome assemblers and then merged their results into single assembly (Sup. Fig. 2). First, IDBA-tran assembler (Peng et al., 2013) was used to directly combine 454 data, non-stranded and stranded Illumina data (although in this mode strand information was not used during the assembly), resulting in 114,859 transcripts (Sup. Fig. 2A). Next, Trinity assembler (Grabherr et al., 2011) was used to assemble kmer-normalized stranded data, resulting in 169,511 transcripts (Sup. Fig. 2B). Furthermore, Mira assembler was used to assemble kmer normalized Illumina data and 454 data, resulting in 130,990 transcripts (Sup. Fig. 2C). Next, the results from these three assemblies were merged step-wise using Minimus2 assembler (Sommer et al., 2007), which resulted in 187,286 transcripts (Sup. Fig. 2D). The assembled transcripts were verified by mapping on them strand-specific Illumina data, and custom scripts were used to remove polymorphisms and indels according to consensus voting. The corrected assembly was further filtered by removing transcripts that matched to longer transcripts with at least 98% of their length and with at least 98% identity, resulting in 174,922 transcripts in the final MLRNA131024 assembly.

The assembly was further annotated in the same way as the previous MLRNA110815 assembly, with transcript clustering, Pfam annotation and matching against human genes.

As shown in Supplementary Figure 2, the N50 size and the length of the longest transcript are large in the combined assembly than in any of the three sub-assemblies, suggesting that merging results of different assemblers might be indeed a productive way for transcriptome reconstruction from current RNA-seq datasets. Furthermore, the number of different Pfam models substantially increased in this assembly from 3,813 to 5,329, and the number of human homologs also increased from 9,151 to 10,268 (Table 2). Manual examination of several transcripts revealed that short indels prevented their correct annotation in Pfam and human homolog searchers in MLRNA110815 assembly (data not shown). Importantly, the rank 4 CEG missing in MLRNA110815 assembly was identified in MLRNA131024 assembly (Table 3). All combined, the new *de novo* transcriptome assembly represents substantial improvement in both accuracy and completeness of the reconstructed transcripts. However, as mentioned previously, the neoblast screen presented in the subsequent sections of this Chapter was initially based on the previous assembly version and for the sake of record we decided not to upgrade this analysis to the newer transcriptome assembly in the scope of this Chapter.

Table 3. Coverage of Core Eukaryotic Genes in *M. lignano* de novo transcriptome assemblies.

	MLRNA110815	**MLRNA131024**
Rank	Present/Total	Present/Total
4	**64**/65	**65**/65
3	61/61	61/61
2	55/56	55/56
1	65/66	65/66
Total	**245**/248 **(98.8%)**	**246**/248 **(99.2%)**

2. Comparison of RNA-seq data from irradiated and intact worms

Comparison between irradiated and non-irradiated transcriptomes has proved to be a successful method for identifying potential neoblast-specific genes in planarian flatworms. Transcripts that are expressed in irradiated neoblast-deficient animals at much lower level than in control population, are likely to be specific for proliferating cells (Rossi et al., 2007; Eisenhoffer et al., 2008; Blythe et al., 2010).

All the proliferating cells of *M. lignano* can be eliminated by fractionated exposure of worms to high doses of γ-irradiation, with no obvious damage to differentiated cells. Worms can live up to 5 weeks after being exposed to lethal doses of irradiation, and during first 10 days they look and behave almost similar to intact worms (De Mulder et al., 2010). First, we prepared and sequenced RNA-seq libraries from the worms 7 days post-irradiation and from intact animals (mock irradiated) to compare transcription profiles of normal and neoblast-deficient populations (Libraries 210G, CTRL1, 210Gi, CTRL1i, Table 1). In this case, the same starting RNA material was used to prepare libraries for both SOLiD and Illumina platforms. The reason for this was that in-house SOLiD system was available to perform all RNA-seq expression

profiling experiments, while some Illumina data were required in order to improve *de novo* transcriptome assembly (construction and sequencing of 210Gi and CTRL1i libraries was outsourced to a commercial provider Baseclear). We also reasoned that comparison of data from different platforms made from the same material might help with filtering out platform-specific biases.

For the initial analysis, RNA-seq reads were mapped to the transcriptome assembly MLRNA110815 and transcript expression was estimated from the mapped reads using RSEM package (Li and Dewey, 2011) and converted to percentage points of the total of mapped reads for comparison between different samples. We focused on transcripts which have at least 0.005% reads count in control samples, log2 value of irradiated/control -0.5 or less, and have identifiable human homologs. This selection resulted in 454 candidate transcripts (Sup. Table 1), and contained such known neoblast marker genes as *Piwi* (Rank 161 by read abundance) and *PCNA* (Rank 388), suggesting that the list should contain true neoblast marker genes. At the same time, a number of house-keeping genes such as tubulin, collagen and some lipoproteins appeared to be downregulated in irradiated worms as well. To assess the quality of the candidate list, we semi-randomly selected 38 transcripts for *in situ* hybridization analysis (Sup. Table 1).

Based on the *in* situ hybridization results we clustered all the observed expression patterns in 5 categories: differentiated tissues, testis, ovaries, both gonads, gonads and neoblasts. A few examples that did not clearly belong to any of these groups we defined as "no clear pattern". ISH results for certain genes representing mentioned categories are presented in Figure 2.

Surprisingly, expression of 5 transcripts out of 38 was detected in differentiated tissues. 7 genes appeared to be testes-specific, 4 – ovary-specific, 11 were expressed in both gonads, for 5 genes we could not obtain detectable expression pattern, and only 7 showed expression in neoblasts and gonads (Fig. 6A). Thus, although we did confirm the expression pattern of several candidate genes to be neoblast-specific, the fraction of such genes in the candidate list was relatively low and the candidate list needed further improvement.

3. Expression profiles of juvenile and adult worms

Since the neoblast gene candidate list obtained by simple comparison of irradiated and non-irradiated animals contained many transcripts with expression in gonads but not necessarily in neoblasts, we decided to refine the candidate list by generating additional data that would allow distinguishing gonad-related related genes. For this, we performed RNA-seq on the worms at 4 days after hatching, and of adult animals, using the SOLiD platform (Libraries Juv and CTRL2, Table 1). At 4 days after hatching worms already have all the organs developed, except for reproductive system. No gonads could be observed yet, though there is already a defined group of cells called gonad anlage that would give rise to germ line (Pfister et al., 2008; De Mulder et al., 2009). Thus, we expect the expression of germ line-specific genes to boost during formation of the gonads. At the same time, upregulation of transcripts expressed mainly in somatic proliferating cells should not be that significant, even if they are expressed in germ cells as well. We used this hypothesis to improve our filtering criteria for neoblast-specific genes and to exclude gonad-specific genes from the candidate list. First, we calculated statistically significant differentially expressed transcripts between irradiated and non-irradiated samples, and between juveniles and adult animals, using NOISeq package (Tarazona et al., 2011), and

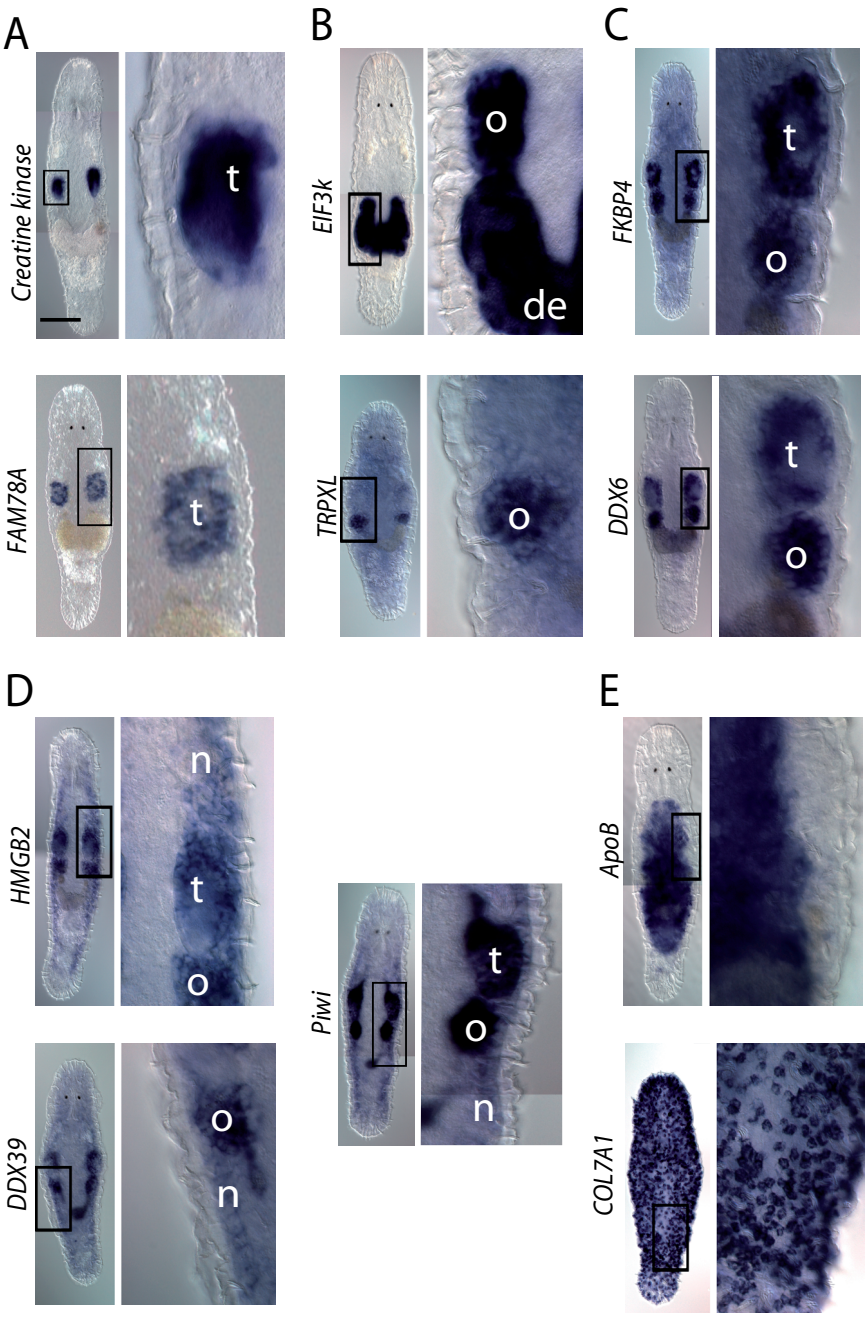


Figure 2. Examples of expression patterns visualized by *in situ* hybridization. **(A)** Genes expressed in testes. **(B)** Genes expressed in ovaries and (sometimes) developing eggs. **(C)** Genes expressed in both gonads. **(D)** Genes expressed in gonads and neoblasts. *Piwi* expression pattern is given as an example of gene expressed in all dividing cells. **(E)** Genes expressed in differentiated tissues. Magnified picture of the area in black frame is shown on the right side of whole-worm picture. t – testis, o – ovary, de – developing egg, n – neoblasts. Scalebar 100 μm (all worms).

next selected for transcripts that are at least 4-fold downregulated after irradiation but not differentially expressed between juveniles and adults or higher in juveniles than in adults.

Discarding the transcripts that do not have clear human homologue, we ended up with 173 candidate neoblast-specific genes (Sup. Table 2). Well-known neoblast marker *Piwi* appeared among these candidates as well (Ranked 141), and generally the list was enriched for cell division factors (such as *cyclins, CDCs* and members of the Ras family), chromatin modifiers (histones and nucleosome assembly factors) and proteins involved in transcription, translation and posttranslational modifications (splicing and polyadenilation factors, ribonucleotide reductases, ribosome proteins and various kinases and phosphatases). We have analyzed expression patterns of 151 candidates from this list by ISH (Sup. Table 2). Only 4 genes of all tested were expressed in differentiated tissues, and expression of 62 was detected in both gonads, and of another 34 – in gonads and neoblasts. Notably, only 4 transcripts appeared to be

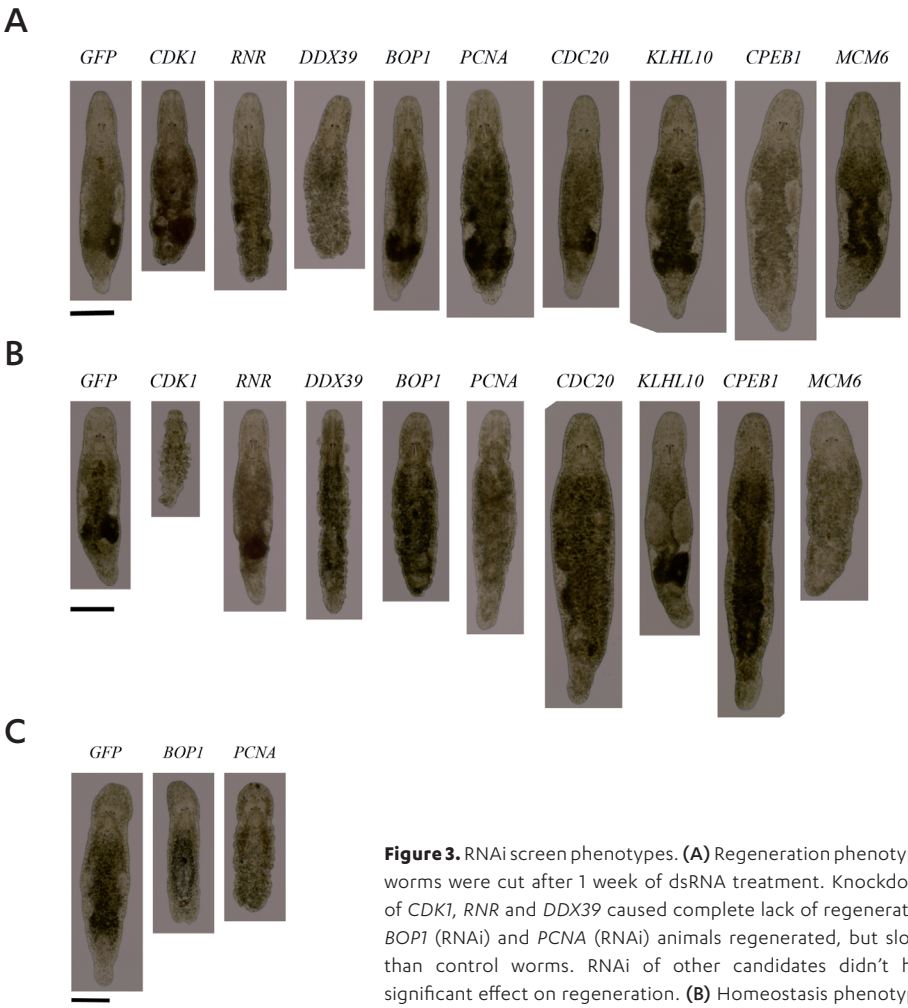


Figure 3. RNAi screen phenotypes. **(A)** Regeneration phenotypes; worms were cut after 1 week of dsRNA treatment. Knockdowns of *CDK1*, *RNR* and *DDX39* caused complete lack of regeneration. *BOP1* (RNAi) and *PCNA* (RNAi) animals regenerated, but slower than control worms. RNAi of other candidates didn't have significant effect on regeneration. **(B)** Homeostasis phenotypes. ▶

ovary-specific, while as many as 38 genes were expressed exclusively in testes. 9 probes did not produce any clear pattern (Fig. 6B). Thus, it appears that filtering out potential gonad-specific genes, as well as applying more stringent statistical approaches (NOISeq) for calculation of differentially expressed genes improves the fraction of neoblast genes in the candidate list.

4. Expression dynamics of genes during the first week after irradiation

In order to study changes in transcription profiles caused by irradiation and consequent elimination of proliferating cells in more detail, we also performed RNA-seq of *M. lignano* worms at 1, 3 and 7 days after irradiation in comparison with intact (mock-irradiated) animals (Libraries D1, D3, D7 and CTRL2, Table 1). Based on these data all transcripts whose expression is affected by irradiation can be clustered in a number of groups. Some transcripts are down-regulated gradually after irradiation, while expression of the others is not affected during few days, and only starts to go down after 3 or even 7 days. There are also a number of genes that get significantly up- or down-regulated right after irradiation, but then their expression slowly gets back to normal.

Each of these groups is likely to represent certain expression pattern and/or function, shared by its members. We focused on the transcripts gradually down-regulated after irradiation, considering them the best candidates to be neoblast-specific. We expect the most important stem cell markers to be also expressed at lower levels in neoblast progeny cells, which do not proliferate and are not therefore affected by irradiation. It takes time in a normal course of differentiation to completely eliminate expression of these genes. Besides, not all of the dividing cells die at once after irradiation. Some of them, probably quiescent and slow-cycling, could still be detected some 5 days after irradiation (De Mulder et al., 2010).

Among all the transcripts that are consistently downregulated during the first week after irradiation (i.e. fold-change increases during the time-course) we selected the ones that have clear human homologue and are expressed at reasonably high levels in control animals (read count >0,002%), ending up with the list of 244 candidate transcripts (Supplementary Table 3). We found many classical cell division and stem cell markers on this list, such as *Piwi* (ranked 24), *PCNA* (ranked 72), different *MCM*s (Salvetti et al., 2000; Orii et al., 2005; De Mulder et al., 2009) and, once again, a large number of genes involved in progression through cell cycle, transcription, translation, posttranscriptional protein modification and chromatin modification.

We checked expression pattern of 65 candidates by ISH (Supplementary Table 3). Remarkably, none of the tested transcripts was expressed in differentiated tissues. 8 genes we defined as testis-specific, 4 – as ovary-specific. Expression of 38 candidates was detected in both gonads, and of another 11 – also in neoblasts. 4 of the probes did not result in any clear pattern (Fig. 6C). Thus, it appears that "consistent decline" criteria yield the most productive list of neoblast candidate genes in our hands.

▶ After 3 weeks of treatment knockdown of *CDK1, RNR, DDX39, BOP1* and *PCNA* caused loss of the gonads and development of bulges. *DDX39* (RNAi) caused the most severe phenotype, and all worms died after 3 weeks, so the pictures were made 2 weeks after the start of the experiment. *CDC20, CBEP1* and *MCM6* knockdowns resulted in complete or partial loss of gonads. *KLHL10* (RNAi) phenotype is characterized by enlarged testes. **(C)** Late regeneration phenotypes. Tail amputation after 3 weeks of *PCNA* (RNAi) and *BOP1* (RNAi) treatment caused complete lack of regeneration. Scalebar 150 μm (all worms).

5. Functional studies of candidate neoblast markers

Of all the transcripts checked in all three approaches, 136 seem specific for proliferating cells as they are expressed in both gonads or in both gonads and neoblasts. These we expect to be the most important ones for stem cell functioning. We randomly picked 47 out of 136 and addressed their function by RNAi experiments, both in intact and regenerating animals. The RNAi screen focused on severe and fast-developing phenotypes, of which 9 were found. These phenotypes could be divided in 3 groups, based on differences observed during regeneration (Fig. 3A). RNAi against three genes (*DDX39, CDK1, RNR*) caused a complete lack of regeneration. In all three cases the wound was closed after cutting but no visible blastema was ever formed. In two cases that we defined as late regenerative phenotypes (*BOP1, PCNA*) blastema was formed but differentiation and growth were delayed. For these genes we performed additional RNAi experiments, with tail amputation performed after three weeks of dsRNA treatment (normally the worms were cut after 1 week of treatment). Late amputation resulted in a complete lack of regeneration (Fig. 3C). RNAi against the last 4 transcripts (*CDC20, KLHL10, CPEB1, MCM6*) caused no obvious regeneration defects. Cut worms regenerated exactly in the same way as *GFP* (RNAi) control animals.

During homeostasis the first two groups developed virtually similar phenotype within 3 weeks, mainly characterized by the loss of gonads and the development of bulges (Fig. 3). The *DDX39* (RNAi) was the only one resulted in the death of all worms within the 3-weeks-screen and quickly developing most severe phenotype. The four phenotypes affecting only homeostasis are limited to the gonads. Knockdown of *CDC20* and *MCM6* resulted in the total

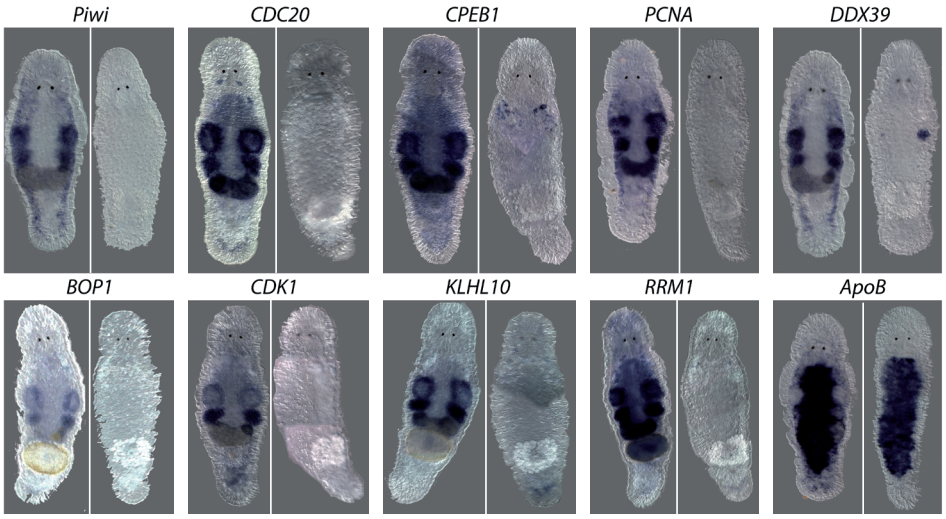


Figure 4. Expression of *Piwi*, *CDC20*, *CPEB1*, *PCNA*, *DDX39*, *BOP1*, *CDK1*, *KLHL10*, *RRM1* and *ApoB* in irradiated animals and under normal conditions. Expression of all the genes except for *ApoB* is restricted to dividing cells and is almost completely gone following irradiation. Gut-specific expression of *ApoB* is not affected by irradiation. More explanation in the text. Scalebar 150 μm (all worms).

loss of gonads, *CPEB1* (RNAi) led to a loss of ovaries, and the *KLHL10* phenotype is characterized by significantly enlarged filled up testes (Fig. 3B).

We then checked whether expression of all the genes that caused these obvious phenotypes was restricted to proliferating cells. In order to do that we compared their expression patterns in intact and irradiated worms 9 days after irradiation, when dividing cells cannot be detected in an animal anymore. *CPEB1* still appeared to be expressed at very low level in gut after irradiation, while the signal in gonads and neoblasts was completely gone. Some trace amounts of *KLHL10* mRNA could also be detected in the differentiated tissues of irradiated worms. Expression of other 6 candidates was restricted to proliferating cells and is not detectable in irradiated worms (Fig. 4).

6. Proliferation rate in knockdown worms

To test whether the observed phenotypes can be related to changes in the proliferation rate, we determined the number of mitotic cells at the 10th day of RNAi treatment in both cut and uncut worms. During homeostasis, a significant ($p < 0.05$) decrease in the number of mitotic cells compared to the *GFP*-controls can be observed during knock down of *DDX39, RNR, CDK1, BOP1, PCNA,* and *MCM6* (Fig. 5).

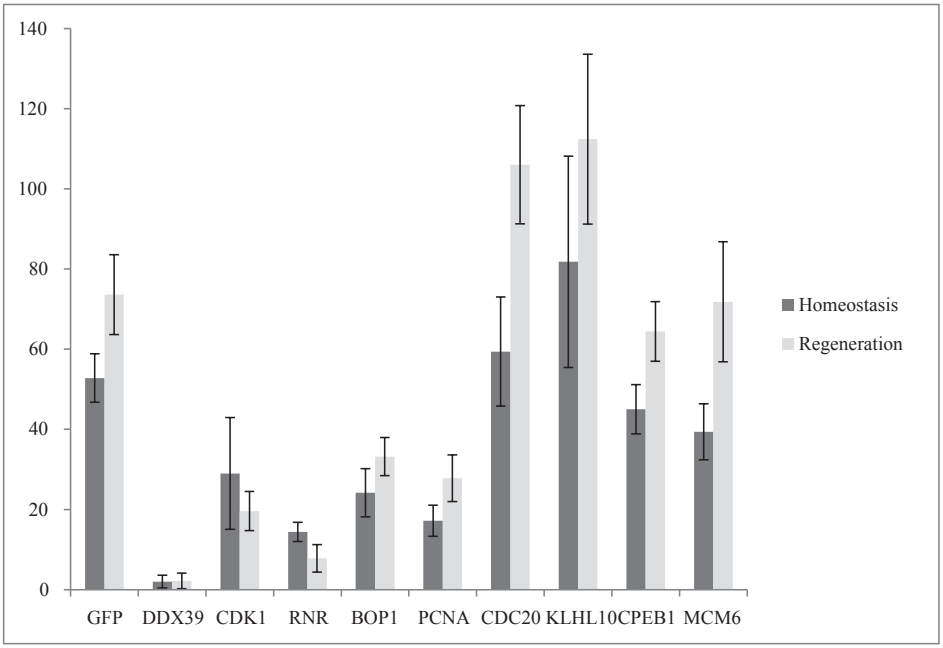


Figure 5. RNAi effect on the number of mitotic cells during regeneration and homeostasis. Knockdown of *DDX39, RNR, CDK1, PCNA, BOP1* and *MCM6* causes significant decrease in the number of mitotic cells during homeostasis. During regeneration *DDX39* (RNAi), *CDK1* (RNAi) and *RNR* (RNAi) animals demonstrated a clear decrease in the number of mitotic cells. More explanation in the text.

It has been shown that regeneration leads to a significant increase in the number of mitotic cells 48 hours after cutting (Nimeth, 2007). Remarkably, during knock down of *DDX39*, *CDK1*, and *RNR*, the number is decreased instead (p=0.862; p=0.215; p=0.009; respectively). In the case of *KLHL10*, there is still an increase, although not significant anymore (p=0.079). All other phenotypes and the *GFP*-control still have a significant increase.

7. DDX39 in human cell lines

DDX39 gene is expressed exclusively in dividing cells, both germ line and somatic, in a pattern similar to stem-cell marker *Piwi* (Sup. Figure 4A). Its expression is also detected in the blastema of regenerating worms (Sup. Figure 4B). RNAi against *DDX39* dramatically affects both tissue homeostasis and regeneration. In order to see whether this function of the gene is conserved we addressed possible roles of human *DDX39* genes (*UAP56* and *URH49* for *DDX39B* and *DDX39A* respectively) in regulation of cell proliferation *in vitro*. We used Hek293 cells to test the effects of overexpression and knockdown of *DDX39* paralogs in cell cultures. Colony formation assay was performed and formed colonies were stained with crystal violate. Both *DDX39A* and *DDX39B* (Sup. Fig. 5 B and E) overexpression resulted in significant increase in proliferation rate in comparison with control (empty vector delivered) cells (Sup. Fig. 5 A and D). Knockdown of *DDX39A* or *DDX39B* (Sup. Fig. 5 C and F) causes drastic decrease of proliferation.

DISCUSSION

1. First assembly of *M. lignano* transcriptome

High-quality annotated transcriptome is a prerequisite for a modern model organism. We have first generated *M. lignano de novo* transcriptome assembly ML110815, which was used as a basis for neoblast marker screen reported in this Chapter, and later further improved it to ML131024 assembly, which was used in another study described in Chapter 5. While *de novo* transcriptome assemblies have their challenges and limitations and usually do not fully cover the actual transcriptome of the organism (Steijger et al., 2013), they are still extremely useful and are often also the only available choice in the absence of full-length cDNA data. The assemblies we present here are rather complete if judged by the coverage of core eukaryotic genes (Table 3) and diversity of present Pfam domains (Supplementary Figures 1 and 2). In the latest assembly we combined several assembly algorithms, which seem to improve the quality of the assembly significantly. At the same time, this assembly includes more than 170,000 transcript, which cover alternatively spliced variants and possibly some underassembled transcripts. Further improvements of the transcriptome assembly will require generation of specific datasets focused on transcript boundaries in order to correctly annotated transcription start sites and 3'UTRs. Genome-assisted assembly would be also very useful in the future in refining the transcriptome.

2. Quantitative RNA-seq approaches allow prediction of gene function and expression pattern

We wanted to check whether the set of transcripts in our assembly is useful for differential gene expression studies with. We used different sets of RNA-seq data to generate list of candidate genes whose expression is potentially enriched in neoblasts.

Rapid elimination of proliferative compartment of flatworms by irradiation has long been used as experimental approach to find neoblast-specific transcripts. Comparison between irradiated and intact samples helped to design riboprobes for localization of pluripotent cells (Shibata et al., 1999; Salvetti et al., 2000), provided gating criteria for sorting neoblast populations (Hayashi et al., 2006) and was successfully used to define radiation-sensitive gene expression profiles (Rossi et al., 2007; Eisenhoffer et al., 2008; Friedländer et al., 2009; Solana et al., 2012; Wagner et al., 2012). We used the same approach to predict a number of genes that are likely to be expressed in dividing cells, removed by irradiation. Expression of the most of the candidate genes we have checked by ISH is restricted to proliferative compartments of the worm body, proving this approach to be a useful tool for prediction of transcripts specific for dividing cells.

Comparison of stage-specific transcriptomes could help to identify the genes important at different time points of development. We tried to use RNA-seq data from juvenile and adult worms to eliminate gonad-specific transcripts from our candidate list and enrich it for transcripts important for stem cell functioning in general. Cutoffs we applied did not improve the prediction accuracy as much as we anticipated (we discuss possible reasons for that below) but still we believe that stage-specific transcription profiles would be very useful for gonad-specific gene prediction and detailed studies of *M. lignano* embryonic and post-embryonic development.

Analyzing gene expression levels at different time points after irradiation helps to understand expression dynamics of possible neoblast-specific genes. It could also lead to discovery of molecular markers for different subpopulations of proliferating cells within an organism. Based on the generated RNA-seq data we can speculate that genes up- or down-regulated immediately after irradiation and then getting to their normal expression levels are likely to be involved in stress-response mechanisms. Transcripts sustaining their expression level for a few days after irradiation and only then being downregulated we would expect to be expressed either in some sort of quiescent slow-cycling stem cells, or in progeny of pluripotent stem cells. Finally, genes showing consistent decline of expression over a time course after irradiation are likely to be specific for pluripotent stem cells, both in germ line and in neoblasts. We checked this last hypothesis and showed that expression of 94% of the genes we predicted to be specific for dividing cells is restricted to proliferative organs of *M. lignano*. Moreover, 75% of all tested transcripts are expressed both in ovaries and testis (all germ cells) or even in both gonads and neoblasts, indicating that function of these genes is important for all populations of proliferating cells.

Combining different sets of RNA-seq data one could better understand functions and importance of given gene and define better criteria to look for genes specific for certain cell population. Adding more data (for example, from starved or stressed animals, regenerating worms at different time points of regeneration or sorted cell populations) would greatly improve our knowledge of gene expression dynamics in *M. lignano* and provide a useful tool for studying regeneration, development and stem cell functioning in general. Our lab is working on creating an online open source database where we plan to collect all possible information about *M. lignano* transcripts, including all kinds of RNA-seq data, expression patterns and RNAi phenotypes.

3. Consistent decline of expression after irradiation is a reliable indicator of stem-cell related genes

In this study we used three alternative approaches to look for genes possibly expressed in pluripotent cells. We generated three different candidate lists, and the last one, based on consistent decline of expression after irradiation proved to be the most reliable.

We checked expression patterns of the candidates from all three lists by ISH and defined genes expressed in differentiated tissues as false positives. Transcripts found exclusively in testes or ovaries are also unlikely to be involved in general stem cell functions due to their organ-specificity. So we only considered genes expressed at least in both gonads to be positive as potential pluripotent cell markers. Somatic pluripotent cells represent only 6,5% of all cells in *M. lignano*. They are also small and are spread in the mesoderm of the worm (Pfister et al., 2007), and that makes it complicated to detect neoblast-specific ISH signal, especially of the transcripts expressed at low levels. Also in all our experiments we never saw a transcript expressed in neoblasts that would not give an ISH signal in both gonads at the same time. So the difference between two groups of expression patterns that we defined in the beginning of this study – namely both gonads and both gonads plus neoblasts, – is rather faint and depends in a lot of cases on the eye of the person analyzing ISH data. Whenever a gene appears to be expressed in both gonads, we cannot rule out the possibility of its presence in neoblasts as well, just below the detectable level. That is the main reason why in this screen we consider all the transcripts found to be expressed in both gonads to be potentially positive also in neoblasts regardless of whether we can see the signal in neoblasts or not.

First list was based purely on comparison of RNA-seq data from irradiated and intact worms, and 15% of transcripts from the list that we tested by ISH appeared to be expressed in differentiated tissues. Another 33% were expressed either in testes, or in ovaries, therefore not being good candidates as general "stemness" markers (Fig. 6A). Such a high percentage of false positives can be explained by the large impact of non-specific effects introduced by whole organism irradiation. Indeed, exposure to lethal doses of γ-irradiation triggers crucial shifts in expression levels of large number of genes, involved in stress response, DNA repair mechanisms, apoptosis etc. These processes occur virtually in every cell of the organism, so one would expect down-regulation of genes involved in related pathways to be even more significant than of neoblast-specific transcripts, expressed only in a small population of cells.

In the second approach we tried to use stage-specific RNA-seq data to get rid of gonad-specific genes. A lot of classical proliferation markers (such as *PCNA* and *MCM*s) were not present in the resulting candidate list. We conclude that transcripts that are important for neoblast proliferation, play even more significant role in germ cells, where proliferation rate is generally higher. This goes in accordance with the fact that all genes we observed to be expressed in neoblasts, always give even stronger ISH signal in gonads. By getting rid of gonad-specific genes we lose all the candidates that have important function in neoblasts as well. Thus by applying stage-specific cutoff we do not enrich the list for neoblast-specific genes, but rather create a selection of genes expressed both in gonads and somatic stem cells, but at rather low levels.

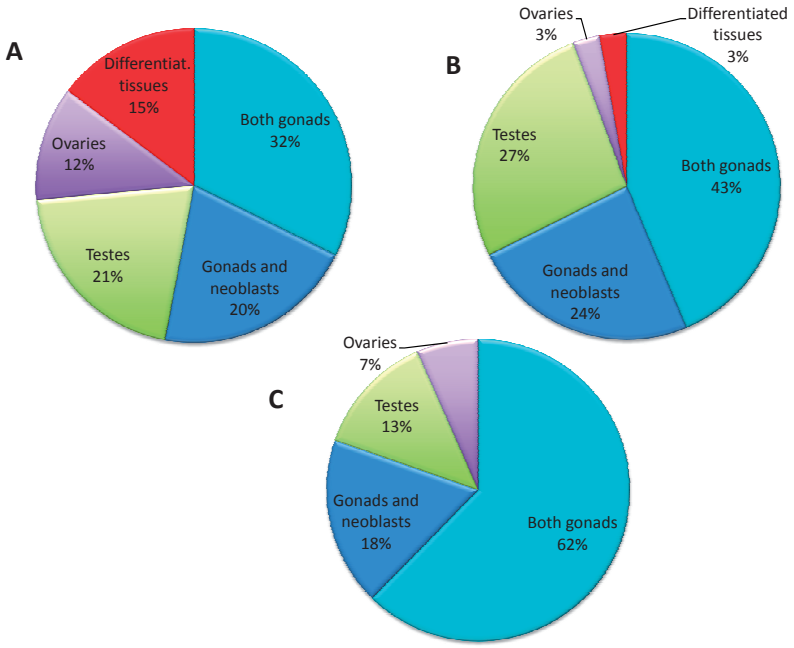


Figure 6. Expression patterns of genes from different candidate lists. **(A)** Expression of 34 candidate genes from the list based on comparison of RNA-seq data sets from irradiated and control (non-irradiated) worms. **(B)** Expression patterns of 142 genes selected as neoblast-specific candidates based on their expression levels in juvenile and adult worms. **(C)** Expression patterns of 61 genes from the list based on consistent decline of expression during the first week after irradiation.

Another interesting result from this list is the abundance of testes-specific genes (27% of tested transcripts). Testis is normally formed earlier during development than ovaries, and already in juveniles groups of cells could be found, called gonad anlage, that would later give rise to testes (Pfister et al., 2008; De Mulder et al., 2009). Consequently, some early testis-specific genes are already expressed in juveniles, and, unlike ovary-specific genes, are not dramatically upregulated during juvenile-to-adult transition. That explains why so many genes on this list are expressed in testes, and only 3% - in ovaries. (Fig. 6B)

Third list provides the best tool out of three for identification of genes whose expression is potentially enriched in proliferating cells. Almost all the candidates from that list that we checked by ISH were expressed in proliferating compartments of the worm body, and 80% are expressed at least in both gonads. Studying the expression dynamics after irradiation gives a good insight into possible function of given gene, and helps to rule out a lot of false positive candidates (that appeared on the first list, for example). Indeed, by picking only the transcripts that are gradually downregulated after irradiation, we are likely to avoid the ones involved in stress response, DNA repair and also the ones expressed in the progeny of pluripotent stem cells. This set of RNA-seq data can be of great use for further studies, since one can find molecular markers for different subpopulations of neoblasts and germ cells by using different

selection criteria. Expanding this data set (performing RNA-seq at more time points after irradiation) can greatly increase accuracy of such search and ultimately give us an insight into heterogeneity of neoblast populations in flatworms.

4. Virtually all the transcripts expressed in neoblasts are found in germ line as well

95% of all the candidate transcripts we checked by ISH were expressed in germ cells – some in testes, some in ovaries, most of the candidates in both gonads. At the same time throughout all our experiments we were unable to identify a single gene expressed exclusively in somatic pluripotent cells, and all the genes we found active in neoblasts were at the same time expressed at high levels in gonads. This corresponds with the data from planarian worms, where many genes expressed in the neoblast population encode proteins similar to the ones found in germ cells of other organisms (Shibata et al., 2010; Wang et al., 2010; Wagner et al., 2012; Solana et al., 2012).

We know that somatic stem cells in *M. lignano* can give origin to germ line (Egger et al., 2006), and it is not completely clear whether the opposite is possible. Relations between neoblasts and germ cells remain largely unknown – we do not know which cell line is evolutionary older, and whether they maintain pluripotency using the same mechanisms. Same question applies to other types of pluripotent cells such as induced pluripotent stem (iPS) cells, embryonic stem (ES) cells and tumor stem cells. There is a large overlap between the sets of genes expressed in neoblasts and germ cells (Shibata et al., 2010; Wang et al., 2010), germ line specific factors are used to create iPS cells (Takahashi and Yamanaka, 2006) and are activated in various types of cancers (De Smet and Loriot, 2013), similar sets of genes are reported to be active in planarian neoblasts and vertebrate stem cells (Onal et al., 2012). Yet we know almost nothing about actual mechanisms that establish and maintain pluripotency in all of these cell types.

M. lignano provides an exceptional model system, where two lineages of pluripotent stem cells could be studied side-by-side and their embryonic origins could be investigated.

5. Most of the knockdowns do not cause any obvious phenotypes

In this study we generated knockdowns of 47 different genes that we expect to be important for stem cell functioning. Surprisingly, less than 20% of them caused any obvious phenotypes in *M. lignano*. Poor performance of the method itself could be almost ruled out as a reason since we checked the efficiency of RNAi by ISH and confirmed that target gene expression is normally completely gone after 2,5 weeks (Sup Figure 1). Still incomplete knockdown might be a reason for a lack of phenotype in a few cases.

Another possible reasons would be functional redundancy of some genes (as in case with *MCM*s, for example) and stability of the proteins encoded by genes we were trying to knockdown. We think it is the experiment setup that mainly explains such a small number of scored phenotypes. In this RNAi screen we only focused on the most severe quickly developing changes in morphology or behavior of the worms. We only followed the worms treated with dsRNA for 3 weeks, and almost never longer. RNAi for *PCNA* and *BOP1* were the only ones we performed for longer, and treated worms only developed strong phenotype 4-5 weeks after the beginning of the experiment. Another indication that we have missed a lot of slowly developing

phenotypes is represented by mitotic labelings of worms. For example, in cases of *PCNA, BOP1* and *MCM6* knockdowns we see proliferative system of the worm being already affected by knockdown as early as after 10 days of treatment, while morphologically no differences with control worms could be detected at this moment. Longer RNAi screens should be considered for better understanding of other candidates' functions.

6. DDX39 is a novel conserved stem cell specific gene

DDX39 knockdown caused the most severe phenotype among all we observed during that screen. The effect of its loss developed even faster than that of *Piwi* (De Mulder et al., 2009), leading to the death of treated worms 2 weeks after the start of the experiment. Its expression can clearly be detected in both gonads and neoblasts and is restricted to proliferating cells. RNAi of *DDX39* in regenerating worms blocks the regeneration completely. We conclude that *DDX39* is a novel conserved stem cell specific gene.

DDX39 is a member of the DEAD box RNA helicase family, which includes multiple members with a common D-E-A-D (Asp-Glu-Ala-Asp) motive. There are highly conserved members of this family in different species, and they are known for their roles in RNA metabolism, ranging from RNA biogenesis to RNA decay (Linder and Jankowsky, 2011). Functions of *DDX39* have been linked to mRNA export in *Drosophila* (Eberl et al., 1997) and *Caenorhabditis elegans* (MacMorris et al., 2003). In adult *Drosophila*, the gene homolog is expressed in follicle and nurse cells (Eberl et al., 1997), and has a role in transporting mRNA from nurse cells to the growing oocyte (Meignin and Davis, 2008) and in scaffolding of the nuage structure and piRNA processing in germ cells (Zhang et al., 2012).

Function of its human homologue has not been studied until 5 years ago, when Sugiura and colleagues first showed that *DDX39* gene is upregulated in lung squamous cell cancer and that its overexpression can stimulate cell growth *in vitro* (Sugiura et al., 2007; Sugiura et al., 2007). Since then *DDX39* was also described as a potential oncogene in different types of cancer (Kubota et al., 2012; Kikuta et al., 2012) and its role in telomere protection and elongation was demonstrated (Yoo and Chung, 2011). In mammals, two paralogs of *DDX39* exist, *DDX39A* and *DDX39B*, also known as *UAP56* and *URH49* respectively. *M. lignano DDX39* homolog is closer to the human *DDX39B/UAP56* gene (Sup. Fig. 6) by several aminoacids.

Overexpression of *DDX39A* or *DDX39B* paralogs leads to remarkable increase in proliferation rate, while knockdown of any of *DDX39* genes results in diminished proliferation *in vitro*, suggesting conservative role of *DDX39* in stem cell regulation. Consistent with these findings are recent studies, that link DDX39 upregulation with tumor growth (Sugiura et al., 2007) and poor prognosis in thyroid carcinoma (Montero-Conde et al., 2008).

CONCLUSIONS

We sequenced the transcriptome of *M. lignano* and completed its first annotated assembly. We then used it as a reference for different sets of quantitative RNA-seq data, which we collected from irradiated worms at different time points after irradiation, worms at different developmental stages and intact control worms. We conclude that the generated *de novo* transcriptome assembly is suitable and useful for these RNA-seq approaches.

Based on RNA-seq data we prepared 3 alternative candidate lists of genes with expression potentially enriched in all proliferating cells and checked actual expression patterns of candidates from every list by ISH. We found consistent decline of expression over a time course after irradiation to be the most reliable selection criterion for potential neoblast-specific genes.

We addressed functions of randomly selected candidates from all 3 lists by RNAi and described 8 stem cell associated knockdown phenotypes, 3 of which had an impact only on homeostasis, and the other 5 affected both homeostasis and regeneration, being likely important for neoblast functions. We describe DDX39 as a novel highly conserved stem cell specific gene, playing a crucial role in cell proliferation.

We conclude that *M. lignano* can be successfully used as a model for stem cell and regeneration research and identification of conserved genes involved in different aspects of stem cell functioning.

MATERIALS AND METHODS

1. Animal culture

M. lignano worms of inbred DV1 line (Janicke et al., 2013) were cultured in Petri dishes in f/2 medium, nutrient-enriched artificial sea water (Andersen et al., 2005), and were fed with unicellular diatom *Nitzschia curvilineata* (Heterokontophyta, Bacillariophyceae), which covers the bottom of the petri dishes (Ladurner et al., 2005). All animals were kept in the incubators under following conditions: 20°C temperature, 60% humidity and 14/10 h day/night cycle (Rieger et al., 1988). For RNAi interference experiments we also used transgenic GFP-expressing Hub1 line (Demircan et al, manuscript in preparation). Culturing conditions for this line were exactly the same, as for DV1 line.

2. Collecting material for RNA-seq libraries

Worms were starved for 18–24 hours prior to RNA isolation to prevent diatom RNA contamination, then rinsed in fresh medium. Total RNA was extracted using TRI Reagent (T9424, Sigma), according to manufacturer's instructions. Animals were homogenized in TRI Reagent by pipetting. For every extraction a batch of 200-300 worms was used. Samples were resuspended in nuclease-free water and treated with 5 U of DNAse I (Thermo Scientific, Cat. No. EN0521) for 45 minutes at 37°C. Enzyme and all the remaining DNA were removed by extraction with phenol : chloroform : isoamyl alcohol (125:24:1, pH 4,5 Life technologies, Cat. No. AM9720). Samples were alcohol precipitated overnight at −80°C. Total RNA was pelleted by centrifugation at 12,000g for 20 mins at 4°C, washed with 70% ethanol and air-dried for 5 minutes. RNA was resuspended in nuclease-free water. Concentration of total RNA samples was measured with Qubit RNA BR assay kit (Invitrogen, Cat. No. Q10211).

For making stage-specific RNA-seq libraries synchronized populations of *M. lignano* were created. 2000-3000 well-fed animals were kept on a single Petri dish for no longer than 2 hours. Then all the worms were moved to different plates, while the eggs that were laid during these two hours stayed on the original dish. Usually around 200-300 synchronized eggs could be found on the plate after this procedure. Juvenile worms were collected for RNA isolation 1 week

after the eggs were synchronized (3-4 days after hatching). Total RNA was isolated from adult population 3 weeks after synchronization (2,5 weeks after hatching).

For RNA-seq analysis of irradiated worms (also at different time points after irradiation) and non-irradiated control animals mixed populations were used.

3. Irradiation by γ-rays

Worms were irradiated by three single successive doses (210 Gy, 15 Gy and 15 Gy) in accordance with "fractionated irradiation" protocol described before (De Mulder et al., 2010). Control (mock-irradiated) animals were handled in the same way as treated worms, only irradiation steps were omitted.

4. Preparation and sequencing of SOLiD RNA-seq libraries

Total RNA for SOLiD libraries was enriched for mRNA using the Poly A Purist Kit

(Life Technologies, Cat. No. AM1919). All SOLiD libraries were made with SOLiD total RNA-seq kit (Life Technologies, Cat. No. 4445374) in accordance with manufacturer's protocol (Publication Part Number 4452437 REV. B; Revision date July 2011) and sequenced on SOLiD 3 platform.

5. Preparation and sequencing of 454 transcriptome libraries

Random-primed normalized cDNA library for 454 sequencing was prepared by Vertis Biotechnologie AG (Freising, Germany). Total RNA was isolated from the worms pellet using the mirVana miRNA isolation kit (Ambion). The RNA preparation was analyzed for its integrity by capillary electrophoresis. From the total RNA poly(A)+ RNA was prepared. First-strand cDNA synthesis was primed with a N6 randomized primer. Then 454 adapters A and B were ligated to the 5' and 3' ends of the cDNA. The cDNA was finally amplified with PCR (16 cycles) using a proof reading enzyme. Normalization was carried out by one cycle of denaturation and reassociation of the cDNA, resulting in N1-cDNA. Reassociated ds-cDNA was separated from the remaining ss-cDNA (normalized cDNA) by passing the mixture over a hydroxylapatite column. After hydroxylapatite chromatography, the ss-cDNA was amplified with 10 PCR cycles. For Titanium sequencing the cDNA in the size range of 500 – 700 bp was eluted from a preparative agarose gel. An aliquot of the size fractionated cDNA was analyzed by capillary electrophoresis. The library was sequenced on GS FLX Titanium machine following manufacturer's protocol.

6. Preparation and sequencing of Illumina RNA-seq libraries

Illumina RNA-seq libraries were prepared by Baseclear BV (Leiden, The Netherlands) using the standard Illumina mRNA-seq sample preparation pipeline and sequenced on Illumina Genome Analyzer IIx using paired-end 100-cycle run.

7. Computational analysis of RNA-seq data

Generation and annotation of *de novo* transcriptome assemblies is described in detail in the results section of this Chapter.

For differential gene expression analysis reads were mapped to MLRNA131024 transcriptome assembly using bowtie v.1.0.0 (Langmead et al., 2009) with the following parameters:

'-q -n 2 -e 99999999 -l 16 -a -m 200 --best –strata'. The mapping results were filtered to exclude hits with more than 5 mismatches (NM:I tag in SAM output) and transcript quantification was performed by RSEM package v.1.2.7 (Li and Dewey, 2011), which takes into account read mapping ambiguity when estimating gene expression values. To calculate transcript fold changes between different samples, the RSEM values were next either converted to read percentages and fold changes calculated directly, or NOIseq software (Tarazona et al., 2011) was used to calculate statistically significant differences.

8. Whole-mount in situ hybridization

Whole mount *in situ* hybridization (ISH) was carried out by following an earlier described protocol (Pfister et al., 2007) and using young adult worms 4-8 weeks old adults. For *Piwi* and *DDX39* ISH on juveniles, animals 5 days after hatching were used. For ISH on regenerating worms young adult animals (4-8 weeks old) were cut above (anterior from) the gonads and fixed 12 hours after cutting. Pictures were made using Nomarski microscope DIC optics and AxioCam HRC (Zeiss) digital camera. Some pictures were stitched and background was adjusted (brightness and contrast) using Adobe Photoshop software.

cDNA synthesis was performed using the SuperScript™ III First-Strand Synthesis System (Life Technologies, Cat. No. 18080-051) following the manufacturer's protocol with 2-3 ug of total RNA as template per reaction. Provided oligo(dT) and hexamer random primers were used.

DNA fragments selected as templates for ISH probes were amplified from cDNA by standard PCR with GoTaq Flexi DNA Polymerase (Promega, Cat. No. M8305), then cloned using pGEM-T vector system (Promega, Cat. No. A3600) and sequenced by in-house facility. All the primers used to amplify these fragments are listed in Supplementary Table 4. DNA templates for producing DIG-labeled riboprobes were amplified from sequenced plasmids using High Fidelity Pfu polymerase (Thermo Scientific, Cat. No. EP0572). Primers used were 5'-CGGCCGCCATGGCCGCGGGA-3' as forward and 5'-TGCAGGCGGCCGCACTAGTG-3' as reverse binding pGEM-T vector backbone right outside the insertion site. Of both primers we also had a version with T7 promoter sequence (5'-GGATCCTAATACGACTCACTATAGG-3') appended upstream to the 5' end to serve as the start site in subsequent *in vitro* transcriptions. A pair of primers, forward with T7 promoter and reverse without or vice versa was used to amplify every ISH probe template. Which pair was used depended on the orientation of the cloned fragment in pGEM-T vector.

Digoxigenin (DIG) labelled RNA probes (500 to 800 bp in length) were generated using the DIG RNA labeling Mix (Roche, Cat. No. 11277073910) and T7 RNA polymerase (Promega, Cat. No. P2075) following manufacturer's protocol for *in vitro* transcription. Concentration of every probe was measured with Qubit RNA BR assay kit (Invitrogen, Cat. No. Q10211), then probes were diluted in Hybridization Mix (Pfister et al., 2007) to 20 ng/µl, stored at -80°C and used within 4 months. Final concentration of the probe and optimal temperature for hybridization varied for different probes and were determined in trials. Starting conditions were 1ng/µl and 55°C.

9. RNA interference (RNAi)

Specific knockdown of candidate genes by RNA interference with double-stranded RNA delivered by soaking was achieved as previously described (De Mulder et al., 2009; Pfister et al., 2008). RNAi soaking experiments were performed on young adult animals (4-8 weeks old at

the start of treatment) in 24-well plates. 15 individuals were used per well and kept on algae in 300 µl of dsRNA solution (30 ng/µl in f/2). dsRNA solution (or f/2 in control experiments) was refreshed daily in every well. Every week all the worms were transferred to fresh 24-well plate with algae. In experiments addressing regeneration worms were cut after 1 week of treatment. All the animals were treated with dsRNA for 3 weeks, then experiments were stopped (except for *PCNA* RNAi treatment that went on for 6 weeks). Individuals used for making pictures and performing mitotic labelings were chosen randomly.

To exclude possible off-target effect of delivered dsRNA, control worms were treated with GFP dsRNA or just f/2 medium (mock). No morphological difference was observed between mock and GFP dsRNA treated control DV1 worms. We also used GFP dsRNA on transgenic GFP-expressing worms of Hub1 line (Demircan et al, manuscript in preparation) as positive control. GFP could no longer be observed in dsRNA-treated animals after 9-14 days of treatment (Sup. Fig. 3A).

To generate dsRNA fragments same plasmids were used as for making ISH probes. Templates for the synthesis of both RNA strands were amplified from the plasmid containing the fragment of interest. Same primers were used as for ISH riboprobe template amplification, and for each fragment two PCRs were performed – with both pairs of primers (forward with T7 promoter sequence upstream and reverse without it and vice versa). Reaction volume was 150 µl for every PCR and High Fidelity Pfu polymerase (Thermo Scientific, Cat. No. EP0572) was used. PCR products were run on 1% agarose gel, the PCR product bands were cut and purified using QIAquick Gel Extraction kit (QIAGEN, Cat. No. 2704). Each template was then used to synthesize the corresponding single strand RNA with TranscriptAid T7 High Yield Transcription Kit (Thermo Scientific, Cat. No. K0441) according to manufacturer's protocol. Single reaction volume was 50 µl, and reactions were incubated at 37°C for 5 hours. Then 75 µl of nuclease-free water was added to every reaction, sense and antisense RNA strands were mixed to a final volume of 300 µl and annealed by incubation at 70°C for 10 minutes and gradual cooling down to room temperature, that took 90 minutes. Every sample was then treated with 1 U of RNAse A (Life Technologies, Cat. No. 12091-039) and 5 U of DNAse I (Thermo Scientific, Cat. No. EN0521) for 45 minutes at 37°C. Samples were alcohol precipitated overnight at −80°C. dsRNA was pelleted by centrifugation at 12,000g for 15 mins at 4°C, washed with 75% ethanol and air-dried for 5 minutes. dsRNA was resuspended in nuclease-free water and concentration was measured using Nanodrop ND1000. Freshly autoclaved and filtered f/2 medium was used to adjust the concentration to 30 ng/µl, samples were aliquoted in 1,5 ml Eppendorf tubes (in a way to use 1 tube a day and thus avoid freezing and thawing) and stored at -80°C.

10. Mitotic labeling

Mitosis labeling was performed as described in (Ladurner et al., 2000). In short, worms were washed in f/2 medium and relaxed in 1:1 $MgCl_2$:f/2 for 5 min, fixed in 4% paraformaldehyde (PFA) for 1h, washed with PBS-T (PBS plus 0.1% Triton X-100) and blocked with BSA-T (1% bovine serum albumin in PBS-T) for 30 min. The primary anti-phospho histone H3 Antibody (Millipore) was diluted 1:100 in BSA-T and applied overnight at 4˚C, followed by washing with PBS-T. Worms were incubated with the secondary Goat anti-Rabbit IgG Antibody (Millipore), diluted 1:150 in BSA-T, for 1h. After being washed with PBS-T, slides were mounted using Vectashield.

11. Cell culture, transfection and colony formation assay

Hek293 cell line used in this study was maintained in DMEM (Sigma, Cat. No. 51444C) supplemented with 10% heat-inactivated fetal bovine serum at 37°C. For over-expression experiments, transient transfection of plasmids OCAAo5051B0937D from Life Sciences for DDX39A and SC11060 from Origene for DDX39B (0.8 g/ml) was performed using 5 µl/ml of Lipofectamine 2000 reagent (Life Technologies, Cat. No. 11668019). Transfections were performed according to manufacturer's instructions. For knockdown of DDX39A or DDX39B, Hek293 cells were transfected with small interfering RNA pool (SMARTpool: ON-TARGETplus DDX39A siRNA L-004920-01-0010 and SMARTpool: ON-TARGETplus DDX39B siRNA L-003805-00-0010 available from ThermoScientific) at final concentration of 100 nM. Transfections were made using Dhermafect 1 (Thermo Scientific, Cat. No. T-2001-03) according to manufacturer's protocol.

48h after transfection cells were trypsinized and counted. 10000 cells for each treatment were seeded in 10-cm plates in the presence of G418 (600 mg/L) and allowed for colony formation. The medium was exchanged every two or three days. After 10 days, the surviving colonies were fixed in MeOH followed by staining with 0.2% crystal violet. The experiment was performed in triplicate.

REFERENCES

Adamidi, C., Wang, Y., Gruen, D., Mastrobuoni, G., You, X., Tolle, D., Dodt, M., Mackowiak, S.D., Gogol-Doering, A., et al. (2011). De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics. Genome Res *21*, 1193-1200.

Adell, T., Salò, E., Boutros, M., and Bartscherer, K. (2009). Smed-Evi/Wntless is required for beta-catenin-dependent and -independent processes during planarian regeneration. Development *136*, 905-910.

Agata, K., Nakajima, E., Funayama, N., Shibata, N., Saito, Y., and Umesono, Y. (2006). Two different evolutionary origins of stem cell systems and their molecular basis. Semin Cell Dev Biol *17*, 503-09.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Andersen, R.A., Berges, R.A., Harrison, P.J., and Watanabe, M.M. (2005). Recipes for freshwater and seawater media; enriched natural seawater media. In: Anderson, R.A. (Ed.), Algal culturing techniques, Elsevier academic press, Burlington, 507.

Baguñà, J. (2012). The planarian neoblast: the rambling history of its origin and some current black boxes. Int J Dev Biol *56*, 19-37.

Blythe, M.J., Kao, D., Malla, S., Rowsell, J., Wilson, R., Evans, D., Jowett, J., Hall, A., Lemay, V., et al. (2010). A dual platform approach to transcript discovery for the planarian Schmidtea mediterranea to establish RNAseq for stem cell and regeneration biology. PLoS One *5*, e15617.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E., Wetter, T., and Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res *14*, 1147-159.

Eberl, D.F., Lorenz, L.J., Melnick, M.B., Sood, V., Lasko, P., and Perrimon, N. (1997). A new enhancer of position-effect variegation in Drosophila melanogaster encodes a putative RNA helicase that binds chromosomes and is regulated by the cell cycle. Genetics *146*, 951-963.

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

Eisenhoffer, G.T., Kang, H., and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in

the planarian Schmidtea mediterranea. Cell Stem Cell *3*, 327-339.

Felix, D.A., and Aboobaker, A.A. (2010). The TALE class homeobox gene Smed-prep defines the anterior compartment for head regeneration. PLoS Genet *6*, e1000915.

Fernandéz-Taboada, E., Moritz, S., Zeuschner, D., Stehling, M., Schöler, H.R., Saló, E., and Gentile, L. (2010). Smed-SmB, a member of the LSm protein superfamily, is essential for chromatoid body organization and planarian stem cell proliferation. Development *137*, 1055-065.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res *39*, W29-W37.

Friedländer, M.R., Adamidi, C., Han, T., Lebedeva, S., Isenbarger, T.A., Hirst, M., Marra, M., Nusbaum, C., Lee, W.L., et al. (2009). High-resolution profiling and discovery of planarian small RNAs. Proc Natl Acad Sci U S A *106*, 11546-551.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics *28*, 3150-52.

Gentile, L., Cebrià, F., and Bartscherer, K. (2011). The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration. Dis Model Mech *4*, 12-19.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol *29*, 644-652

Guo, T., Peters, A.H., and Newmark, P.A. (2006). A Bruno-like gene is required for stem cell maintenance in planarians. Dev Cell *11*, 159-169.

Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008). Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science *319*, 323-27.

Hayashi, T., and Agata, K. (2012). A unique FACS method to isolate stem cells in planarian. Methods Mol Biol *879*, 29-37.

Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Dev Growth Differ *48*, 371-380.

Huang, X., and Madan, A. (1999). CAP3: A DNA sequence assembly program. Genome Res *9*, 868-877

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D.B., and Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution *67*, 3233-242.

Kikuta, K., Kubota, D., Saito, T., Orita, H., Yoshida, A., Tsuda, H., Suehara, Y., Katai, H., Shimada, Y., et al. (2012). Clinical proteomics identified ATP-dependent RNA helicase DDX39 as a novel biomarker to predict poor prognosis of patients with gastrointestinal stromal tumor. J Proteomics *75*, 1089-098.

Kubota, D., Okubo, T., Saito, T., Suehara, Y., Yoshida, A., Kikuta, K., Tsuda, H., Katai, H., Shimada, Y., et al. (2012). Validation study on pfetin and ATP-dependent RNA helicase DDX39 as prognostic biomarkers in gastrointestinal stromal tumour. Jpn J Clin Oncol *42*, 730-741.

Labbé, R.M., Irimia, M., Currie, K.W., Lin, A., Zhu, S.J., Brown, D.D., Ross, E.J., Voisin, V., Bader, G.D., et al. (2012). A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. Stem Cells *30*, 1734-745.

Ladurner, P., Rieger, R., Baguñà, J. (2000). Spatial distribution and differentiation potential of stem cells in hatchlings and adults in the marine platyhelminth Macrostomum sp.: a bromodeoxyuridine analysis. Dev Biol *226*, 231-241.

Ladurner, P., Pfister, D., Seifarth, C., Schärer, L., Mahlknecht, M., Salvenmoser, W., Gerth, R., Marx, F., and Rieger, R. (2005). Production and characterisation of cell- and tissue-specific monoclonal antibodies for the flatworm Macrostomum sp. Histochem Cell Biol *123*, 89-104.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Langmead, B.,Trapnell, C., Pop, M., and Salzberg, S.L.(2009). Ultrafastand memory-efficient alignmentof short DNA sequences to the human genome. Genome Biol *10*, R25.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Lin, H.T., Otsu, M., and Nakauchi, H. (2013). Stem cell therapy: an exercise in patience and prudence. Philos Trans R Soc Lond B Biol Sci *368*, 20110334.

Linder, P., and Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. Nat Rev Mol Cell Biol *12*, 505-516.

MacMorris, M., Brocker, C., and Blumenthal, T. (2003). UAP56 levels affect viability and mRNA export in Caenorhabditis elegans. RNA *9*, 847-857.

Meignin, C., and Davis, I. (2008). UAP56 RNA helicase is required for axis specification and cytoplasmic mRNA localization in Drosophila. Dev Biol *315*, 89-98.

Montero-Conde, C., Martín-Campos, J.M., Lerma, E., Gimenez, G., Martínez-Guitarte, J.L., Combalía, N., Montaner, D., Matías-Guiu, X., Dopazo, J., et al.

(2008). Molecular profiling related to poor prognosis in thyroid carcinoma. Combining gene expression data and biological information. Oncogene *27*, 1554-561.

Morgan, T.H. (1898). Experimental Studies of the Regeneration of Planaria Maculata. Arch. Entwicklungsmech. Org. *7*, 364-397.

Morgan, T.H. (1902). Growth and regeneration in Planaria lugubris. Arch. Ent. Mech. Org. 13.

Morris, J., Ladurner, P., Rieger, R., Pfister, D., Del Mar De Miguel-Bonet, M., Jacobs, D., and Hartenstein, V. (2006). The Macrostomum lignano EST database as a molecular resource for studying platyhelminth development and phylogeny. Dev Genes Evol *216*, 695-707.

Morris, J., Nallur, R., Ladurner, P., Egger, B., Rieger, R., and Hartenstein, V. (2004). The embryonic development of the flatworm Macrostomum sp. Dev Genes Evol *214*, 220-239.

De Mulder, K., Kuales, G., Pfister, D., Egger, B., Seppi, T., Eichberger, P., Borgonie, G., and Ladurner, P. (2010). Potential of Macrostomum lignano to recover from gamma-ray irradiation. Cell Tissue Res *339*, 527-542.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., et al. (2009). Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. Dev Biol *334*, 198-212.

Newmark, P.A., and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev Biol *220*, 142-153.

Newmark, P.A., Reddien, P.W., Cebrià, F., and Sánchez Alvarado, A. (2003). Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. Proc Natl Acad Sci U S A *100 Suppl 1*, 11861-65.

Nishimura, O., Hirao, Y., Tarui, H., and Agata, K. (2012). Comparative transcriptome analysis between planarian Dugesia japonica and other platyhelminth species. BMC Genomics *13*, 289.

Okano, H., Nakamura, M., Yoshida, K., Okada, Y., Tsuji, O., Nori, S., Ikeda, E., Yamanaka, S., and Miura, K. (2013). Steps toward safe cell therapy using induced pluripotent stem cells. Circ Res *112*, 523-533.

Onal, P., Grün, D., Adamidi, C., Rybak, A., Solana, J., Mastrobuoni, G., Wang, Y., Rahn, H.P., Chen, W., et al. (2012). Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. EMBO J *31*, 2755-769.

Orii, H., Sakurai, T., and Watanabe, K. (2005). Distribution of the stem cells (neoblasts) in the planarian Dugesia japonica. Dev Genes Evol *215*, 143-157.

Oviedo, N.J., Nicolas, C.L., Adams, D.S., and Levin, M. (2008). Live Imaging of Planarian Membrane Potential Using DiBAC4(3). CSH Protoc *2008*, pdb.prot5055.

Parra, G., Bradnam, K., Ning, Z., Keane, T., and Korf, I. (2009). Assessing the gene space in draft genomes. Nucleic Acids Res *37*, 289-297.

Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., and Sánchez Alvarado, A. (2009). Formaldehyde-based whole-mount in situ hybridization method for planarians. Dev Dyn *238*, 443-450.

Peng, Y., Leung, H.C., Yiu, S.M., Lv, M.J., Zhu, X.G., and Chin, F.Y. (2013). IDBA-tran: a more robust de novo de Bruijn graph assembler for transcriptomes with uneven expression levels. Bioinformatics *29*, i326-334.

Petersen, C.P., and Reddien, P.W. (2008). Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. Science *319*, 327-330.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. (2007). The exceptional stem cell system of Macrostomum lignano: screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. Front Zool *4*, 9.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., et al. (2012). The Pfam protein families database. Nucleic Acids Res *40*, D290-D301.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science *310*, 1327-330.

Rieger, R.M., Gehlen, M., Haszprunar, G., Holmlund, M., Legniti, A., Salvenmoser, W., and Tyler, S. (1988). Laboratory cultures of marine Macrostomida (Turbellaria). Fortschr Zool *36*, 523.

Rink, J.C. (2013). Stem cell systems and regeneration in planaria. Dev Genes Evol *223*, 67-84.

Rink, J.C., Gurley, K.A., Elliott, S.A., and Sánchez Alvarado, A. (2009). Planarian Hh signaling regulates regeneration polarity and links Hh pathway evolution to cilia. Science *326*, 1406-410.

Robb, S.M., Ross, E., and Sánchez Alvarado, A. (2008). SmedGD: the Schmidtea mediterranea genome database. Nucleic Acids Res *36*, D599-D606.

Rossi, L., Salvetti, A., Marincola, F.M., Lena, A., Deri, P., Mannini, L., Batistoni, R., Wang, E., and Gremigni, V. (2007). Deciphering the molecular machinery of stem cells: a look at the neoblast gene expression profile. Genome Biol *8*, R62.

Rouhana, L., Shibata, N., Nishimura, O., and Agata, K. (2010). Different requirements for conserved post-transcriptional regulators in planarian regeneration and stem cell maintenance. Dev Biol *341*, 429-443.

Salvetti, A., Rossi, L., Deri, P., and Batistoni, R. (2000). An MCM2-related gene is expressed in proliferating cells of intact and regenerating planarians. Dev Dyn *218*, 603-614.

Sandmann, T., Vogg, M.C., Owlarn, S., Boutros, M., and Bartscherer, K. (2011). The head-regeneration transcriptome of the planarian Schmidtea mediterranea. Genome Biol *12*, R76.

Sánchez Alvarado, A. (2004). Regeneration and the need for simpler model organisms. Philos Trans R Soc Lond B Biol Sci *359*, 759-763.

Sánchez Alvarado, A., and Newmark, P.A. (1999). Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proc Natl Acad Sci U S A *96*, 5049-054.

Scimone, M.L., Meisel, J., and Reddien, P.W. (2010). The Mi-2-like Smed-CHD4 gene is required for stem cell differentiation in the planarian Schmidtea mediterranea. Development *137*, 1231-241.

Shibata, N., Hayashi, T., Fukumura, R., Fujii, J., Kudome-Takamatsu, T., Nishimura, O., Sano, S., Son, F., Suzuki, N., et al. (2012). Comprehensive gene expression analyses in pluripotent stem cells of a planarian, Dugesia japonica. Int J Dev Biol *56*, 93-102.

Shibata, N., Rouhana, L., and Agata, K. (2010). Cellular and molecular dissection of pluripotent adult somatic stem cells in planarians. Dev Growth Differ *52*, 27-41.

Shibata, N., Umesono, Y., Orii, H., Sakurai, T., Watanabe, K., and Agata, K. (1999). Expression of vasa(vas)-related genes in germline cells and totipotent somatic stem cells of planarians. Dev Biol *206*, 73-87.

De Smet, C., and Loriot, A. (2013). DNA hypomethylation and activation of germline-specific genes in cancer. Adv Exp Med Biol *754*, 149-166.

Solana, J., Kao, D., Mihaylova, Y., Jaber-Hijazi, F., Malla, S., Wilson, R., and Aboobaker, A. (2012). Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNA-seq, RNA interference and irradiation approach. Genome biology *13*, R19.

Sommer, D.D., Delcher, A.L., Salzberg, S.L., and Pop, M. (2007). Minimus: a fast, lightweight genome assembler. BMC Bioinformatics *8*, 64.

Steijger, T., Abril, J.F., Engström, P.G., Kokocinski, F., Abril, J.F., Akerman, M., Alioto, T., Ambrosini, G., Antonarakis, S.E., et al. (2013). Assessment of transcript reconstruction methods for RNA-seq. Nat Methods *10*, 1177-184.

Sugiura, T., Nagano, Y., and Noguchi, Y. (2007). DDX39, upregulated in lung squamous cell cancer, displays RNA helicase activities and promotes cancer cell growth. Cancer Biol Ther *6*, 957-964.

Sugiura, T., Sakurai, K., and Nagano, Y. (2007). Intracellular characterization of DDX39, a novel growth-associated RNA helicase. Exp Cell Res *313*, 782-790.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. (2011). Differential expression in RNA-seq: a matter of depth. Genome Res *21*, 2213-223.

Umesono, Y., Watanabe, K., and Agata, K. (1997). A planarian orthopedia homolog is specifically expressed in the branch region of both the mature and regenerating brain. Dev Growth Differ *39*, 723-27.

Wagner, D.E., Ho, J.J., and Reddien, P.W. (2012). Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. Cell Stem Cell *10*, 299-311.

Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. Science *332*, 811-16.

Wang, Y., Stary, J.M., Wilhelm, J.E., and Newmark, P.A. (2010). A functional genomic screen in planarians identifies novel regulators of germ cell development. Genes Dev *24*, 2081-092.

Weissman, I.L. (2000). Stem cells: units of development, units of regeneration, and units in evolution. Cell *100*, 157-168.

Yoo, H.H., and Chung, I.K. (2011). Requirement of DDX39 DEAD box RNA helicase for genome integrity and telomere protection. Aging Cell *10*, 557-571.

Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N., Vreven, T., Meignin, C., Davis, I., et al. (2012). UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell *151*, 871-884.

SUPPLEMENTARY

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization.

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
1	RNA815_123.3	0.082	0.255	0.068	0.367	0.367	-1.630	-2.427	APOB	gut
2	RNA815_7478.2	0.060	0.116	0.083	0.152	0.152	-0.955	-0.870	PPIF	Neoblasts, gonads
3	RNA815_9915.1	0.081	0.123	0.076	0.135	0.135	-0.609	-0.833	RPL3	
4	RNA815_1155.1	0.027	0.085	0.035	0.126	0.126	-1.680	-1.843	DDX6	gonads
5	RNA815_13313	0.021	0.041	0.028	0.108	0.108	-0.975	-1.931	TUBA1A	
6	RNA815_17001	0.036	0.070	0.057	0.105	0.105	-0.946	-0.891	RPS3A	
7	RNA815_6242.1	0.002	0.075	0.003	0.100	0.100	-5.331	-5.024	TUBA1B	testes
8	RNA815_19729.1	0.031	0.050	0.044	0.098	0.098	-0.683	-1.146	CABP7	
9	RNA815_4929.1	0.035	0.050	0.055	0.095	0.095	-0.520	-0.782	ACTB	
10	RNA815_9188.1	0.047	0.068	0.052	0.077	0.077	-0.511	-0.577	RP11-864I4.1	
11	RNA815_9604	0.031	0.049	0.047	0.074	0.074	-0.660	-0.655	GNB2L1	
12	RNA815_2682.1	0.044	0.062	0.047	0.071	0.071	-0.505	-0.586	YWHAZ	
13	RNA815_4702.1	0.002	0.054	0.004	0.069	0.069	-4.895	-4.076	TUBB2C	testes
14	RNA815_1310.1	0.002	0.079	0.001	0.064	0.079	-5.555	-5.575	TTLL6	
15	RNA815_18037.1	0.025	0.036	0.024	0.061	0.061	-0.537	-1.322	HUWE1	
16	RNA815_15414	0.036	0.054	0.035	0.056	0.056	-0.591	-0.692	NACA	
17	RNA815_17072	0.042	0.072	0.034	0.054	0.072	-0.774	-0.688	RPL7A	
18	RNA815_27691	0.030	0.052	0.030	0.053	0.053	-0.779	-0.810	RPL21	
19	RNA815_5759	0.023	0.038	0.022	0.053	0.053	-0.722	-1.286	CACNA1B	lateral to gut
20	RNA815_24696	0.040	0.061	0.037	0.052	0.061	-0.598	-0.513	RPL23A	
21	RNA815_15092	0.021	0.033	0.031	0.051	0.051	-0.656	-0.750	PRDX6	
22	RNA815_25604.1	0.027	0.042	0.031	0.050	0.050	-0.628	-0.687	RPS27A	
23	RNA815_324.1	0.039	0.060	0.027	0.050	0.060	-0.633	-0.898	SCUBE2	
24	RNA815_4116.1	0.033	0.054	0.027	0.050	0.054	-0.707	-0.860	GLUL	
25	RNA815_16979	0.036	0.056	0.029	0.049	0.056	-0.625	-0.726	RPL12	
26	RNA815_16738	0.000	0.032	0.000	0.048	0.048	-6.080	-6.982	GLIPR1L1	
27	RNA815_2657	0.017	0.054	0.012	0.046	0.054	-1.688	-1.957	HSPD1	Neoblasts, gonads
28	RNA815_29108.1	0.025	0.043	0.029	0.046	0.046	-0.787	-0.642	RPS15	
29	RNA815_2623.1	0.001	0.033	0.002	0.046	0.046	-4.723	-4.568	TUBA1B	testes
30	RNA815_15540	0.022	0.032	0.023	0.045	0.045	-0.548	-0.979	RPSA	
31	RNA815_4631	0.001	0.032	0.001	0.044	0.044	-5.137	-5.484	ELAVL3	
32	RNA815_9158.1	0.001	0.028	0.001	0.043	0.043	-5.766	-6.141	PRKRA	
33	RNA815_8383.1	0.000	0.027	0.001	0.043	0.043	-5.854	-5.238	CKB	testes
34	RNA815_788.1	0.026	0.044	0.025	0.043	0.044	-0.741	-0.801	INVS	
35	RNA815_32552.1	0.019	0.029	0.028	0.042	0.042	-0.598	-0.566	RPL30	
36	RNA815_27241.1	0.016	0.025	0.025	0.041	0.041	-0.611	-0.689	RPL22	
37	RNA815_2067	0.001	0.038	0.001	0.040	0.040	-5.538	-5.295	TTLL2	
38	RNA815_10403	0.021	0.033	0.019	0.040	0.040	-0.629	-1.044	RPL4	
39	RNA815_35134	0.025	0.044	0.025	0.039	0.044	-0.808	-0.636	RPS25	
40	RNA815_24602	0.015	0.021	0.024	0.039	0.039	-0.514	-0.695	RPL18A	
41	RNA815_11321.1	0.024	0.039	0.024	0.039	0.039	-0.693	-0.728	RPL27A	
42	RNA815_33495	0.033	0.047	0.026	0.039	0.047	-0.512	-0.559	RPL22	
43	RNA815_8425	0.001	0.029	0.002	0.039	0.039	-5.288	-4.617	TUBA1A	
44	RNA815_7117	0.004	0.034	0.004	0.038	0.038	-3.052	-3.137	PTMA	
45	RNA815_46333	0.017	0.024	0.022	0.037	0.037	-0.548	-0.770	RPL38	
46	RNA815_10534.1	0.024	0.045	0.019	0.037	0.045	-0.872	-0.922	RPL5	
47	RNA815_24381	0.017	0.028	0.025	0.036	0.036	-0.689	-0.501	BTF3L4	
48	RNA815_2928	0.025	0.038	0.022	0.036	0.038	-0.637	-0.714	ATP5B	
49	RNA815_5698.1	0.001	0.044	0.001	0.035	0.044	-5.655	-4.745	HSPB6	gonads
50	RNA815_33849	0.016	0.031	0.018	0.034	0.034	-0.948	-0.975	RPS25	
51	RNA815_26783	0.022	0.036	0.019	0.034	0.036	-0.687	-0.884	RPS18	
52	RNA815_19135	0.013	0.020	0.020	0.034	0.034	-0.640	-0.721	CRNN	
53	RNA815_31810	0.010	0.021	0.015	0.034	0.034	-1.071	-1.158	RPLP2	
54	RNA815_28186	0.014	0.024	0.021	0.033	0.033	-0.845	-0.659	RPL35	
55	RNA815_125.1	0.028	0.046	0.021	0.033	0.046	-0.710	-0.642	URGCP	
56	RNA815_14904	0.001	0.030	0.001	0.033	0.033	-5.423	-5.017	DNAJB8	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
57	RNA815_29632	0.010	0.019	0.020	0.033	0.033	-0.841	-0.705	RPS20	
58	RNA815_25310	0.016	0.027	0.017	0.033	0.033	-0.750	-0.924	RPL36A	
59	RNA815_26013	0.023	0.045	0.019	0.033	0.045	-0.941	-0.788	RPL14	
60	RNA815_7008	0.000	0.025	0.001	0.032	0.032	-5.988	-5.334	PPP1CB	
61	RNA815_14000	0.025	0.037	0.022	0.032	0.037	-0.568	-0.504	RPS24	
62	RNA815_25698	0.019	0.031	0.022	0.032	0.032	-0.679	-0.519	RPL26	
63	RNA815_3854	0.011	0.021	0.013	0.032	0.032	-0.950	-1.256	SEC61A2	
64	RNA815_4101.1	0.015	0.036	0.013	0.031	0.036	-1.301	-1.301	CCT5	gonads
65	RNA815_25860.1	0.024	0.041	0.019	0.031	0.041	-0.767	-0.700	RPS23	
66	RNA815_9315	0.001	0.020	0.001	0.031	0.031	-4.652	-5.251	AC010614.2	
67	RNA815_31214	0.015	0.026	0.022	0.031	0.031	-0.785	-0.532	RPS15A	
68	RNA815_4147	0.001	0.019	0.002	0.031	0.031	-4.728	-4.147	TUBB2C	
69	RNA815_35805.1	0.025	0.040	0.021	0.031	0.040	-0.720	-0.551	RPL31	
70	RNA815_9343.1	0.007	0.014	0.016	0.031	0.031	-1.059	-0.916	HMGB2	Neoblasts, gonads
71	RNA815_18310.1	0.017	0.032	0.014	0.031	0.032	-0.913	-1.144	RPS3	
72	RNA815_11501	0.011	0.016	0.020	0.030	0.030	-0.556	-0.614	COX4I1	
73	RNA815_732	0.011	0.022	0.011	0.030	0.030	-1.014	-1.416	CYP3A5	
74	RNA815_22949	0.019	0.029	0.020	0.030	0.030	-0.622	-0.563	RPL13A	
75	RNA815_18150	0.011	0.024	0.010	0.030	0.030	-1.096	-1.612	RPS8	
76	RNA815_30426	0.019	0.033	0.021	0.030	0.033	-0.811	-0.501	RPS15A	
77	RNA815_24419.1	0.019	0.029	0.013	0.030	0.030	-0.623	-1.207	RPL32	
78	RNA815_22534	0.025	0.036	0.020	0.029	0.036	-0.545	-0.577	RPL17	
79	RNA815_39952	0.019	0.034	0.015	0.029	0.034	-0.835	-1.021	RPS21	
80	RNA815_5848.1	0.003	0.019	0.002	0.029	0.029	-2.491	-3.886	TUBA1B	
81	RNA815_21800	0.016	0.022	0.016	0.029	0.029	-0.517	-0.870	RPL18A	
82	RNA815_11320.1	0.022	0.036	0.018	0.029	0.036	-0.711	-0.719	RPL10A	
83	RNA815_20777.1	0.002	0.018	0.003	0.029	0.029	-2.921	-3.286	SCARF1	
84	RNA815_19464	0.022	0.036	0.016	0.029	0.036	-0.677	-0.807	RPS9	
85	RNA815_22376	0.011	0.020	0.015	0.029	0.029	-0.897	-0.964	RPL13A	
86	RNA815_4883	0.009	0.033	0.009	0.029	0.033	-1.869	-1.703	SET	
87	RNA815_9261.1	0.008	0.014	0.013	0.028	0.028	-0.876	-1.145	SERP2	
88	RNA815_4680	0.016	0.027	0.013	0.028	0.028	-0.793	-1.107	EIF3D	
89	RNA815_41172	0.012	0.020	0.018	0.028	0.028	-0.689	-0.639	RPS27L	
90	RNA815_13519.1	0.015	0.024	0.017	0.028	0.028	-0.630	-0.707	RPSA	
91	RNA815_7168.1	0.008	0.014	0.016	0.028	0.028	-0.751	-0.792	DAZAP1	
92	RNA815_27870	0.006	0.012	0.014	0.028	0.028	-1.071	-0.940	SLC25A5	
93	RNA815_2640	0.001	0.019	0.001	0.027	0.027	-4.874	-4.541	CPEB1	
94	RNA815_6480.1	0.000	0.016	0.001	0.027	0.027	-5.964	-5.716	TUBA1B	
95	RNA815_8699	0.013	0.023	0.018	0.027	0.027	-0.816	-0.592	RAN	gonads
96	RNA815_25435	0.012	0.022	0.017	0.027	0.027	-0.874	-0.646	RPLP2	
97	RNA815_6753	0.021	0.043	0.013	0.027	0.043	-1.020	-1.047	DDX39B	Neoblasts, gonads
98	RNA815_5138.1	0.023	0.045	0.016	0.026	0.045	-0.971	-0.704	NAP1L1	
99	RNA815_4415.1	0.001	0.022	0.000	0.026	0.026	-5.020	-6.904	ELAVL4	
100	RNA815_12223.1	0.031	0.053	0.016	0.026	0.053	-0.782	-0.689	RPS2	
101	RNA815_26689	0.012	0.021	0.015	0.026	0.026	-0.830	-0.769	RPS13	
102	RNA815_28607	0.021	0.039	0.015	0.026	0.039	-0.897	-0.832	RPS19	
103	RNA815_8346	0.021	0.041	0.014	0.026	0.041	-0.941	-0.939	ALKBH3	
104	RNA815_21935	0.013	0.023	0.014	0.026	0.026	-0.859	-0.913	RPS7	
105	RNA815_26744.1	0.014	0.024	0.017	0.026	0.026	-0.814	-0.559	RPL26	
106	RNA815_1457	0.001	0.027	0.001	0.026	0.027	-5.623	-4.925	TTLL1	
107	RNA815_14252	0.011	0.018	0.018	0.026	0.026	-0.700	-0.510	RPS26	
108	RNA815_1663.1	0.001	0.015	0.001	0.025	0.025	-4.822	-5.440	WSCD2	
109	RNA815_14338.3	0.000	0.014	0.002	0.025	0.025	-5.562	-3.670	WSCD1	
110	RNA815_7554	0.009	0.021	0.008	0.025	0.025	-1.281	-1.564	PRMT1	
111	RNA815_14658.1	0.005	0.012	0.010	0.025	0.025	-1.191	-1.328	SSR3	
112	RNA815_2882.1	0.011	0.023	0.008	0.025	0.025	-1.118	-1.577	CCT7	
113	RNA815_6519.1	0.001	0.017	0.001	0.025	0.025	-4.562	-4.600	TUBA1B	
114	RNA815_12035.1	0.007	0.055	0.006	0.024	0.055	-2.892	-1.903	RPS17	
115	RNA815_18233.1	0.004	0.007	0.013	0.024	0.024	-0.882	-0.826	RPS18	
116	RNA815_817	0.004	0.017	0.013	0.023	0.023	-2.120	-0.822	ATP1A1	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
117	RNA815_9325	0.012	0.021	0.011	0.023	0.023	-0.790	-1.060	RPL17	
118	RNA815_6861.1	0.010	0.016	0.015	0.023	0.023	-0.664	-0.621	ATP5F1	
119	RNA815_21872	0.010	0.017	0.011	0.023	0.023	-0.729	-1.014	RPS9	
120	RNA815_30684	0.016	0.026	0.012	0.023	0.026	-0.719	-0.909	RPL23	
121	RNA815_2335.1	0.002	0.015	0.001	0.023	0.023	-3.173	-4.517	TRAF4	
122	RNA815_11033.1	0.010	0.023	0.010	0.023	0.023	-1.220	-1.143	RPL7	
123	RNA815_4163.1	0.014	0.027	0.015	0.022	0.027	-0.975	-0.574	SERBP1	
124	RNA815_28136	0.012	0.021	0.015	0.022	0.022	-0.768	-0.546	RPS27L	
125	RNA815_3376	0.001	0.020	0.001	0.022	0.022	-4.073	-4.493	TUBB2C	
126	RNA815_11385	0.005	0.010	0.009	0.022	0.022	-0.979	-1.229	DAP	
127	RNA815_17336.1	0.024	0.042	0.013	0.022	0.042	-0.790	-0.786	RPS6	
128	RNA815_5473	0.012	0.019	0.014	0.022	0.022	-0.597	-0.680	EIF3L	
129	RNA815_4632	0.008	0.015	0.010	0.022	0.022	-0.823	-1.125	HMGB2	Neoblasts, gonads
130	RNA815_5454.1	0.002	0.021	0.001	0.022	0.022	-3.765	-4.726	KPNA7	
131	RNA815_4030	0.004	0.009	0.010	0.021	0.021	-1.125	-1.157	MCAT	
132	RNA815_9086	0.009	0.014	0.012	0.021	0.021	-0.659	-0.800	TMED4	
133	RNA815_12816.1	0.005	0.012	0.009	0.021	0.021	-1.360	-1.201	PRDX1	
134	RNA815_28608.1	0.008	0.012	0.012	0.021	0.021	-0.657	-0.864	RPS14	
135	RNA815_19777.1	0.017	0.028	0.014	0.021	0.028	-0.690	-0.635	RPL10	
136	RNA815_21701.1	0.013	0.019	0.014	0.021	0.021	-0.604	-0.595	RPL13	
137	RNA815_23126	0.000	0.013	0.001	0.021	0.021	-6.801	-5.111	SRRM2	
138	RNA815_1151.1	0.010	0.031	0.006	0.021	0.031	-1.655	-1.806	RP11-88G17.6	
139	RNA815_12337.1	0.000	0.013	0.000	0.021	0.021	-6.471	-8.709	ZFP36L2	
140	RNA815_5735	0.001	0.018	0.001	0.021	0.021	-4.674	-4.777	PPP2CB	
141	RNA815_1244.1	0.008	0.016	0.010	0.021	0.021	-0.963	-1.022	SSR1	
142	RNA815_59	0.001	0.020	0.001	0.021	0.021	-4.179	-4.343	KRTAP10-7	
143	RNA815_4958	0.004	0.012	0.008	0.021	0.021	-1.485	-1.422	ANXA13	
144	RNA815_4117.1	0.000	0.012	0.000	0.020	0.020	-5.246	-7.085	CSNK1A1L	
145	RNA815_27692	0.015	0.028	0.013	0.020	0.028	-0.915	-0.666	RPS16	
146	RNA815_26357	0.012	0.023	0.009	0.020	0.023	-0.953	-1.111	RPL12	ovaries
147	RNA815_30526	0.017	0.028	0.013	0.020	0.028	-0.672	-0.630	RPL35A	
148	RNA815_5605.5	0.000	0.010	0.000	0.020	0.020	-7.398	-6.511	TNRC18	
149	RNA815_8128	0.014	0.025	0.012	0.020	0.025	-0.824	-0.749	PTGES3	
150	RNA815_8363	0.015	0.024	0.012	0.020	0.024	-0.662	-0.764	HNRNPA2B1	
151	RNA815_8400	0.010	0.016	0.013	0.020	0.020	-0.700	-0.625	EIF3M	
152	RNA815_23233	0.000	0.013	0.000	0.020	0.020	-6.617	-6.867	CNFN	
153	RNA815_5674.1	0.016	0.027	0.013	0.020	0.027	-0.721	-0.626	DNAJA1	
154	RNA815_4781	0.003	0.015	0.002	0.020	0.020	-2.157	-3.026	GCDH	
155	RNA815_48147	0.010	0.020	0.011	0.020	0.020	-0.944	-0.878	RPL38	
156	RNA815_27597	0.012	0.024	0.011	0.020	0.024	-0.975	-0.829	RPL28	
157	RNA815_21158.1	0.004	0.012	0.008	0.019	0.019	-1.449	-1.340	MEP1B	
158	RNA815_29383	0.009	0.016	0.011	0.019	0.019	-0.872	-0.850	FAU	
159	RNA815_1298.1	0.009	0.020	0.009	0.019	0.020	-1.113	-1.080	HSP90B1	
160	RNA815_3870	0.000	0.016	0.000	0.019	0.019	-6.415	-6.376	CSNK1E	
161	RNA815_813.1	0.002	0.016	0.001	0.019	0.019	-3.045	-3.840	PIWIL1	
162	RNA815_18520.1	0.009	0.014	0.009	0.019	0.019	-0.595	-1.046	RPS5	
163	RNA815_27192	0.011	0.016	0.013	0.019	0.019	-0.574	-0.510	RPL37A	
164	RNA815_31571	0.010	0.015	0.013	0.019	0.019	-0.623	-0.514	RPS26	
165	RNA815_14128	0.017	0.025	0.013	0.019	0.025	-0.510	-0.533	KHDRBS2	
166	RNA815_108	0.006	0.053	0.002	0.019	0.053	-3.065	-3.288	COL7A1	epidermis
167	RNA815_1978.1	0.001	0.016	0.001	0.019	0.019	-4.682	-4.746	WSCD2	testes
168	RNA815_27772.1	0.007	0.011	0.011	0.018	0.018	-0.632	-0.751	UBA52	
169	RNA815_13520.1	0.000	0.013	0.001	0.018	0.018	-5.128	-4.500	TSPAN18	
170	RNA815_2881	0.007	0.013	0.012	0.018	0.018	-0.857	-0.587	ELOVL6	
171	RNA815_950.1	0.004	0.017	0.003	0.018	0.018	-2.226	-2.802	ZFP36L1	
172	RNA815_1477.1	0.001	0.015	0.001	0.018	0.018	-3.702	-4.917	LRP2	
173	RNA815_5052.1	0.009	0.015	0.007	0.017	0.017	-0.697	-1.291	HNRNPK	
174	RNA815_6209	0.007	0.015	0.011	0.017	0.017	-1.134	-0.635	TUBB2A	
175	RNA815_5193	0.004	0.013	0.004	0.017	0.017	-1.823	-2.180	SUMO3	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
176	RNA815_5768.1	0.007	0.012	0.010	0.017	0.017	-0.801	-0.783	SKP1	
177	RNA815_717.1	0.002	0.017	0.001	0.017	0.017	-2.944	-3.716	CLCA1	testes
178	RNA815_3157.1	0.000	0.025	0.000	0.017	0.025	-5.735	-7.574	TTLL10	
179	RNA815_1323	0.001	0.016	0.001	0.017	0.017	-3.799	-4.528	RRM1	
180	RNA815_13423	0.006	0.010	0.011	0.017	0.017	-0.868	-0.580	FKBP1B	
181	RNA815_682.1	0.011	0.023	0.007	0.017	0.023	-1.030	-1.213	EIF3B	
182	RNA815_14801	0.009	0.013	0.009	0.016	0.016	-0.535	-0.918	RPLP0	
183	RNA815_3930	0.011	0.017	0.008	0.016	0.017	-0.552	-1.065	TCP1	
184	RNA815_11859.1	0.009	0.025	0.006	0.016	0.025	-1.508	-1.426	RPL15	
185	RNA815_2404.1	0.011	0.017	0.008	0.016	0.017	-0.674	-0.940	CCT6A	
186	RNA815_8458	0.007	0.016	0.007	0.016	0.016	-1.220	-1.127	GNB2L1	
187	RNA815_1149	0.003	0.018	0.005	0.016	0.018	-2.637	-1.772	MACROD2	
188	RNA815_3268.1	0.009	0.014	0.008	0.016	0.016	-0.654	-0.910	CCT4	no pattern
189	RNA815_9159	0.008	0.014	0.008	0.016	0.016	-0.840	-0.955	PSMA6	
190	RNA815_9635	0.007	0.012	0.008	0.016	0.016	-0.797	-0.965	TMED10	
191	RNA815_17184.1	0.005	0.008	0.008	0.016	0.016	-0.751	-0.915	SSR4	
192	RNA815_757	0.009	0.014	0.007	0.016	0.016	-0.686	-1.256	STT3B	
193	RNA815_1214.1	0.000	0.013	0.000	0.016	0.016	-5.395	-7.806	RB1CC1	
194	RNA815_6155.1	0.000	0.010	0.000	0.016	0.016	-5.490	-5.480	EHMT1	
195	RNA815_3199	0.001	0.013	0.001	0.016	0.016	-4.568	-4.477	FHL2	
196	RNA815_14838.1	0.005	0.015	0.006	0.016	0.016	-1.605	-1.310	SCARF1	
197	RNA815_2417.1	0.001	0.012	0.001	0.016	0.016	-3.246	-3.888	PRKACA	
198	RNA815_12818	0.000	0.022	0.001	0.015	0.022	-6.201	-4.893	MICALL2	
199	RNA815_12396	0.009	0.014	0.006	0.015	0.015	-0.643	-1.303	EIF3I	
200	RNA815_9461	0.005	0.011	0.008	0.015	0.015	-1.072	-0.998	ATP1B1	
201	RNA815_14563	0.000	0.012	0.000	0.015	0.015	-5.466	-5.877	F11	
202	RNA815_15688	0.004	0.007	0.008	0.015	0.015	-1.016	-0.994	SEC61B	gonads
203	RNA815_4430.1	0.000	0.013	0.000	0.015	0.015	-6.511	-5.499	CSNK1E	
204	RNA815_3154.1	0.001	0.014	0.001	0.015	0.015	-4.134	-4.786	COL6A6	
205	RNA815_5606.1	0.009	0.028	0.005	0.015	0.028	-1.569	-1.536	PURA	gonads
206	RNA815_17923	0.007	0.013	0.008	0.015	0.015	-0.870	-0.796	CALN1	
207	RNA815_1798.1	0.006	0.010	0.008	0.015	0.015	-0.675	-0.896	RPN1	
208	RNA815_675.1	0.001	0.020	0.001	0.015	0.020	-4.375	-3.876	KCNMA1	
209	RNA815_3741.1	0.005	0.011	0.006	0.014	0.014	-1.085	-1.288	CCT3	
210	RNA815_54	0.001	0.030	0.000	0.014	0.030	-4.567	-5.472	KRTAP4-4	
211	RNA815_6280	0.006	0.012	0.008	0.014	0.014	-1.142	-0.843	BTF3	no pattern
212	RNA815_9583.1	0.004	0.008	0.006	0.014	0.014	-1.020	-1.197	PHB2	
213	RNA815_1945.1	0.010	0.021	0.008	0.014	0.021	-1.121	-0.803	CCT2	
214	RNA815_476	0.000	0.012	0.001	0.014	0.014	-4.866	-4.180	COL12A1	
215	RNA815_47849	0.009	0.017	0.009	0.014	0.017	-0.895	-0.551	RPL39	
216	RNA815_3095	0.000	0.010	0.000	0.014	0.014	-5.304	-6.336	CTIF	
217	RNA815_9531	0.011	0.020	0.009	0.014	0.020	-0.940	-0.604	RPL6	
218	RNA815_1039.1	0.005	0.012	0.007	0.013	0.013	-1.149	-1.008	TUBB2C	
219	RNA815_20996	0.009	0.019	0.007	0.013	0.019	-1.023	-0.899	RPS17	
220	RNA815_4558	0.000	0.008	0.000	0.013	0.013	-6.891	-5.734	ACTB	
221	RNA815_18911.1	0.001	0.008	0.003	0.013	0.013	-2.619	-2.350	SMCP	head
222	RNA815_5835	0.001	0.011	0.001	0.013	0.013	-4.238	-3.644	FAM78A	
223	RNA815_28505	0.000	0.006	0.001	0.013	0.013	-5.723	-4.424	GLIPR2	
224	RNA815_8713.1	0.006	0.010	0.008	0.013	0.013	-0.620	-0.763	AGRN	
225	RNA815_8936	0.001	0.016	0.000	0.013	0.016	-4.825	-4.762	CSAD	
226	RNA815_12419	0.005	0.009	0.007	0.013	0.013	-0.894	-0.798	PSMD8	
227	RNA815_2434	0.000	0.013	0.000	0.013	0.013	-4.962	-6.006	CSNK1E	
228	RNA815_2184	0.024	0.061	0.003	0.013	0.061	-1.360	-2.192	BMPER	
229	RNA815_16955.1	0.004	0.009	0.005	0.013	0.013	-1.269	-1.214	H2AFX	gonads
230	RNA815_26878	0.007	0.012	0.008	0.013	0.013	-0.799	-0.693	RPL34	
231	RNA815_9012	0.010	0.015	0.008	0.013	0.015	-0.617	-0.695	PSMA8	
232	RNA815_22452.1	0.012	0.020	0.005	0.013	0.020	-0.751	-1.297	RPS12	no pattern
233	RNA815_5690	0.006	0.011	0.007	0.013	0.013	-0.795	-0.907	TRAM1	
234	RNA815_3838.1	0.000	0.019	0.000	0.013	0.019	-6.497	-7.299	LCP1	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
235	RNA815_2297.1	0.001	0.013	0.001	0.013	0.013	-4.593	-3.807	RRBP1	
236	RNA815_6024	0.008	0.016	0.008	0.012	0.016	-0.887	-0.685	SUCLG2	
237	RNA815_9069	0.000	0.011	0.000	0.012	0.012	-5.740	-8.700	RP11-631M21.2	
238	RNA815_5403	0.001	0.013	0.000	0.012	0.013	-4.603	-5.075	KREMEN1	
239	RNA815_3450.1	0.000	0.007	0.001	0.012	0.012	-4.326	-3.681	MEGF6	
240	RNA815_35386	0.007	0.011	0.007	0.012	0.012	-0.723	-0.771	RPL35A	
241	RNA815_1716.1	0.000	0.011	0.000	0.012	0.012	-5.273	-6.572	TTL	
242	RNA815_8088	0.002	0.009	0.003	0.012	0.012	-2.308	-2.286	H1F0	
243	RNA815_38924	0.005	0.009	0.008	0.012	0.012	-0.880	-0.703	RPL30	
244	RNA815_1880.1	0.009	0.014	0.007	0.012	0.014	-0.622	-0.824	GSPT1	
245	RNA815_2034	0.000	0.017	0.000	0.012	0.017	-5.946	-5.735	TUBA4A	
246	RNA815_10718	0.000	0.011	0.000	0.012	0.012	-5.991	-7.076	FAM78B	
247	RNA815_44402	0.006	0.010	0.007	0.012	0.012	-0.762	-0.784	RPL37	
248	RNA815_2794.1	0.000	0.009	0.000	0.012	0.012	-5.848	-4.881	CSNK1E	
249	RNA815_29733	0.004	0.010	0.006	0.012	0.012	-1.172	-1.057	RPS10	
250	RNA815_9864.1	0.000	0.007	0.000	0.012	0.012	-5.887	-6.650	FAM78B	
251	RNA815_26500.1	0.003	0.010	0.005	0.012	0.012	-1.515	-1.129	HSPE1	
252	RNA815_2817.1	0.006	0.020	0.002	0.012	0.020	-1.828	-2.474	RP11-88G17.6	
253	RNA815_2897	0.001	0.012	0.001	0.012	0.012	-3.427	-4.120	KPNA2	
254	RNA815_14880	0.007	0.012	0.008	0.012	0.012	-0.786	-0.582	EIF3K	ovaries
255	RNA815_1618.1	0.000	0.011	0.000	0.012	0.012	-5.012	-7.387	TNXB	
256	RNA815_4617	0.000	0.011	0.001	0.012	0.012	-4.646	-3.610	TUBA1B	
257	RNA815_4945	0.000	0.007	0.000	0.012	0.012	-5.537	-5.869	CSNK1E	
258	RNA815_6382	0.005	0.007	0.007	0.012	0.012	-0.536	-0.701	PSMA1	
259	RNA815_37485.1	0.005	0.008	0.007	0.012	0.012	-0.627	-0.819	RPL37A	
260	RNA815_1741.1	0.007	0.011	0.006	0.012	0.012	-0.605	-1.005	GFPT2	
261	RNA815_254	0.000	0.016	0.000	0.012	0.016	-5.655	-5.858	ABCA10	
262	RNA815_1957	0.000	0.015	0.000	0.012	0.015	-6.773	-6.858	TTLL10	
263	RNA815_35452	0.009	0.016	0.007	0.012	0.016	-0.839	-0.651	RPL35	
264	RNA815_346.1	0.006	0.010	0.007	0.012	0.012	-0.712	-0.726	RANBP2	
265	RNA815_5205	0.002	0.009	0.002	0.011	0.011	-2.041	-2.581	DKC1	
266	RNA815_25861	0.003	0.007	0.004	0.011	0.011	-1.445	-1.509	HSPE1	
267	RNA815_3582	0.005	0.010	0.004	0.011	0.011	-1.060	-1.486	GPD1	
268	RNA815_705	0.005	0.011	0.004	0.011	0.011	-1.264	-1.529	IFT88	
269	RNA815_5605.1	0.002	0.018	0.008	0.011	0.018	-3.007	-0.577	MT-CO2	
270	RNA815_2874	0.007	0.010	0.008	0.011	0.011	-0.611	-0.591	UGCG	
271	RNA815_9615	0.004	0.011	0.005	0.011	0.011	-1.608	-1.108	PTGES3	
272	RNA815_4365	0.000	0.011	0.000	0.011	0.011	-5.207	-6.569	CSNK1A1L	
273	RNA815_1350.1	0.007	0.011	0.007	0.011	0.011	-0.578	-0.729	GALNT2	
274	RNA815_8989	0.000	0.011	0.001	0.011	0.011	-4.536	-4.288	CSAD	
275	RNA815_2450.1	0.000	0.006	0.000	0.011	0.011	-4.359	-6.823	BOLL	
276	RNA815_17185.1	0.004	0.008	0.006	0.011	0.011	-1.088	-0.958	FTH1	
277	RNA815_3807	0.000	0.008	0.000	0.011	0.011	-5.112	-5.257	CSNK1G3	
278	RNA815_4845	0.001	0.009	0.001	0.011	0.011	-3.407	-3.095	SLC18A3	
279	RNA815_3206.1	0.004	0.050	0.001	0.011	0.050	-3.832	-3.712	ABCF2	
280	RNA815_4067.1	0.000	0.011	0.000	0.011	0.011	-5.252	-4.640	CSNK1E	
281	RNA815_369.1	0.007	0.010	0.007	0.011	0.011	-0.504	-0.604	UNC13B	
282	RNA815_7784.1	0.010	0.014	0.006	0.011	0.014	-0.512	-0.872	PSMA5	
283	RNA815_677.1	0.005	0.011	0.006	0.011	0.011	-0.950	-0.777	LRRC71	
284	RNA815_14561	0.003	0.006	0.006	0.011	0.011	-0.954	-0.900	SEC61G	
285	RNA815_5034	0.004	0.013	0.005	0.011	0.013	-1.765	-1.192	FKBP4	gonads
286	RNA815_2145	0.000	0.014	0.000	0.010	0.014	-5.837	-5.028	TTLL4	
287	RNA815_4172.1	0.004	0.008	0.005	0.010	0.010	-0.935	-0.931	TARDBP	
288	RNA815_6795	0.004	0.008	0.006	0.010	0.010	-1.018	-0.919	RBBP4	
289	RNA815_1690	0.009	0.013	0.005	0.010	0.013	-0.557	-1.022	PSMD2	
290	RNA815_9022.1	0.004	0.009	0.006	0.010	0.010	-1.061	-0.740	HMGB2	Neoblasts, gonads
291	RNA815_3185	0.005	0.008	0.005	0.010	0.010	-0.711	-1.084	SLC3A2	
292	RNA815_3435	0.001	0.010	0.000	0.010	0.010	-4.289	-4.647	ACCN1	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
293	RNA815_2968	0.006	0.009	0.005	0.010	0.010	-0.693	-0.915	CLPTM1	
294	RNA815_8508.1	0.000	0.008	0.000	0.010	0.010	-6.361	-8.402	PPP1CC	
295	RNA815_9797	0.009	0.015	0.007	0.010	0.015	-0.677	-0.535	PSMA4	
296	RNA815_14195	0.004	0.006	0.007	0.010	0.010	-0.551	-0.511	OSTC	
297	RNA815_12778	0.000	0.011	0.000	0.010	0.011	-5.266	-6.810	PPPDE1	
298	RNA815_7325	0.007	0.013	0.006	0.010	0.013	-0.780	-0.849	OLA1	
299	RNA815_44995.1	0.006	0.011	0.006	0.010	0.011	-0.857	-0.623	RPS28	
300	RNA815_1138.1	0.003	0.014	0.002	0.010	0.014	-2.239	-2.555	ACSL1	
301	RNA815_7123	0.003	0.005	0.005	0.010	0.010	-0.750	-0.967	SGTA	
302	RNA815_4445	0.004	0.008	0.005	0.010	0.010	-1.173	-0.936	NCL	
303	RNA815_6282	0.006	0.010	0.006	0.010	0.010	-0.836	-0.754	HELT	
304	RNA815_12722	0.001	0.007	0.001	0.010	0.010	-3.711	-3.493	CAV3	
305	RNA815_13458	0.006	0.010	0.006	0.010	0.010	-0.849	-0.703	EIF3G	
306	RNA815_11581	0.008	0.013	0.006	0.009	0.013	-0.736	-0.700	CCT8	
307	RNA815_4851	0.003	0.010	0.003	0.009	0.010	-1.569	-1.765	PAICS	
308	RNA815_3489	0.000	0.006	0.000	0.009	0.009	-4.250	-4.694	WSCD2	testes
309	RNA815_24456	0.002	0.007	0.003	0.009	0.009	-1.770	-1.580	RFC1	
310	RNA815_18913	0.003	0.005	0.005	0.009	0.009	-0.918	-0.951	NDUFAB1	
311	RNA815_21199	0.000	0.009	0.000	0.009	0.009	-5.068	-6.536	ELAVL4	
312	RNA815_1403	0.002	0.007	0.001	0.009	0.009	-2.051	-2.919	NF1	
313	RNA815_3344	0.006	0.009	0.006	0.009	0.009	-0.573	-0.697	STT3A	
314	RNA815_3701	0.009	0.014	0.006	0.009	0.014	-0.518	-0.510	HSPA5	
315	RNA815_22841	0.009	0.013	0.005	0.009	0.013	-0.510	-0.902	RPL24	
316	RNA815_10419	0.006	0.008	0.006	0.009	0.009	-0.539	-0.645	ADK	
317	RNA815_6907	0.005	0.008	0.005	0.009	0.009	-0.707	-0.884	TUFM	
318	RNA815_4884.1	0.002	0.005	0.004	0.009	0.009	-1.190	-1.005	FAH	
319	RNA815_5143.1	0.001	0.008	0.001	0.009	0.009	-3.380	-4.139	PSMC4	
320	RNA815_13698.3	0.001	0.009	0.002	0.009	0.009	-2.784	-1.998	FTH1	gut
321	RNA815_2829	0.004	0.008	0.003	0.009	0.009	-1.235	-1.453	RPN2	
322	RNA815_5712.1	0.001	0.008	0.001	0.009	0.009	-2.978	-3.117	CCNB1	
323	RNA815_5742	0.006	0.011	0.004	0.009	0.011	-0.972	-1.081	CSNK2A1	
324	RNA815_19363	0.000	0.006	0.001	0.009	0.009	-4.769	-4.040	WSCD2	
325	RNA815_11799	0.004	0.007	0.004	0.009	0.009	-0.892	-1.193	SSB	
326	RNA815_1734	0.002	0.009	0.002	0.009	0.009	-1.899	-1.866	ARIH1	
327	RNA815_3335	0.003	0.025	0.001	0.009	0.025	-3.009	-3.858	NFXL1	
328	RNA815_1344.1	0.003	0.008	0.004	0.009	0.009	-1.258	-1.131	FUBP3	
329	RNA815_2329.1	0.004	0.006	0.005	0.009	0.009	-0.613	-0.883	LASS5	
330	RNA815_2207.1	0.006	0.010	0.005	0.009	0.010	-0.605	-0.704	DPP3	
331	RNA815_8818	0.006	0.011	0.006	0.009	0.011	-0.822	-0.617	EIF3H	
332	RNA815_10975	0.003	0.008	0.004	0.009	0.009	-1.495	-1.053	CBX1	
333	RNA815_16244	0.000	0.006	0.000	0.009	0.009	-5.306	-4.312	WSCD2	
334	RNA815_2571.1	0.002	0.007	0.002	0.009	0.009	-2.152	-2.498	MFSD2B	gonads
335	RNA815_5483	0.003	0.005	0.003	0.008	0.008	-1.052	-1.730	TCTEX1D1	
336	RNA815_2299	0.006	0.008	0.004	0.008	0.008	-0.551	-1.024	GYS1	
337	RNA815_4125.1	0.001	0.008	0.001	0.008	0.008	-2.709	-2.842	CCDC55	
338	RNA815_24606	0.000	0.019	0.000	0.008	0.019	-6.209	-5.623	CALM3	
339	RNA815_5133	0.004	0.006	0.005	0.008	0.008	-0.532	-0.768	TMEM120B	
340	RNA815_1003.1	0.007	0.011	0.003	0.008	0.011	-0.696	-1.374	TARS	
341	RNA815_30943	0.005	0.012	0.004	0.008	0.012	-1.288	-1.133	RPS24	
342	RNA815_44996	0.006	0.010	0.004	0.008	0.010	-0.881	-1.065	RPS28	
343	RNA815_8130.1	0.001	0.005	0.001	0.008	0.008	-2.604	-2.538	TBCEL	
344	RNA815_4437.1	0.006	0.014	0.004	0.008	0.014	-1.202	-0.907	NCL	
345	RNA815_6019.1	0.006	0.009	0.004	0.008	0.009	-0.670	-1.047	SARS	
346	RNA815_1184	0.000	0.009	0.000	0.008	0.009	-6.164	-6.485	IDE	
347	RNA815_3675	0.001	0.011	0.001	0.008	0.011	-3.964	-2.893	CSNK1E	
348	RNA815_6256.1	0.005	0.008	0.005	0.008	0.008	-0.546	-0.632	PSMC4	
349	RNA815_5162.1	0.004	0.009	0.004	0.008	0.009	-1.100	-0.844	NOP56	
350	RNA815_2074	0.002	0.005	0.003	0.008	0.008	-1.444	-1.532	MYO1C	
351	RNA815_7885.1	0.002	0.006	0.003	0.008	0.008	-1.493	-1.546	GLIPR2	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
352	RNA815_1176	0.005	0.007	0.005	0.008	0.008	-0.560	-0.544	SQSTM1	
353	RNA815_4315.1	0.000	0.006	0.000	0.008	0.008	-6.547	-5.916	PPEF1	
354	RNA815_12577	0.002	0.006	0.004	0.008	0.008	-1.682	-0.996	TUSC3	gonads
355	RNA815_700.1	0.001	0.005	0.002	0.008	0.008	-2.607	-2.266	MEX3B	ovaries
356	RNA815_3856.1	0.001	0.009	0.000	0.008	0.009	-3.689	-4.609	DLK2	
357	RNA815_3228	0.000	0.007	0.000	0.008	0.008	-4.179	-4.480	MEGF8	
358	RNA815_850.1	0.004	0.008	0.003	0.008	0.008	-0.961	-1.342	KPNB1	
359	RNA815_8006	0.000	0.007	0.000	0.008	0.008	-8.379	-7.282	CAPN11	
360	RNA815_4156	0.005	0.008	0.003	0.008	0.008	-0.694	-1.337	ABCE1	
361	RNA815_8409.1	0.003	0.007	0.005	0.008	0.008	-1.087	-0.504	IDH3B	
362	RNA815_9606	0.001	0.005	0.001	0.008	0.008	-2.649	-3.229	IDH2	
363	RNA815_3299	0.005	0.009	0.004	0.008	0.009	-0.870	-1.116	TARDBP	
364	RNA815_2839.1	0.006	0.009	0.004	0.008	0.009	-0.671	-0.802	EIF2A	
365	RNA815_1931.1	0.004	0.006	0.005	0.008	0.008	-0.586	-0.691	AKAP1	
366	RNA815_6702	0.005	0.009	0.004	0.007	0.009	-0.747	-0.782	HDAC2	
367	RNA815_2021	0.002	0.013	0.003	0.007	0.013	-2.366	-1.439	WASL	
368	RNA815_9651.1	0.004	0.007	0.005	0.007	0.007	-0.812	-0.676	EIF3F	
369	RNA815_946	0.004	0.010	0.002	0.007	0.010	-1.254	-1.905	MTHFD1	
370	RNA815_2559.1	0.001	0.006	0.001	0.007	0.007	-2.875	-3.214	PDE7A	
371	RNA815_863	0.003	0.009	0.003	0.007	0.009	-1.457	-1.536	DNAJC7	
372	RNA815_1600.1	0.004	0.010	0.002	0.007	0.010	-1.179	-2.032	TUBA1C	
373	RNA815_2759	0.003	0.007	0.001	0.007	0.007	-1.307	-2.367	ASAH2	
374	RNA815_273.1	0.008	0.013	0.005	0.007	0.013	-0.757	-0.582	TOP2B	no pattern
375	RNA815_2749	0.005	0.008	0.004	0.007	0.008	-0.736	-0.771	EPRS	
376	RNA815_7489.1	0.001	0.006	0.001	0.007	0.007	-2.672	-2.563	CDK1	
377	RNA815_1521	0.000	0.008	0.000	0.007	0.008	-4.863	-4.326	DNAJB7	
378	RNA815_310.1	0.005	0.009	0.004	0.007	0.009	-0.670	-0.975	TDRD9	
379	RNA815_21801.1	0.003	0.005	0.003	0.007	0.007	-0.912	-1.297	ERH	
380	RNA815_14267.1	0.003	0.006	0.002	0.007	0.007	-0.905	-1.510	PHB	
381	RNA815_4346	0.000	0.005	0.000	0.007	0.007	-4.968	-5.526	CSNK1E	
382	RNA815_1050.1	0.006	0.009	0.005	0.007	0.009	-0.567	-0.515	IMMT	
383	RNA815_4355.1	0.004	0.006	0.004	0.007	0.007	-0.611	-0.666	PCYT2	
384	RNA815_4250	0.000	0.007	0.000	0.007	0.007	-5.728	-7.843	PRKACG	
385	RNA815_7900	0.004	0.006	0.004	0.007	0.007	-0.658	-0.917	PLD3	
386	RNA815_4953.1	0.005	0.009	0.004	0.007	0.009	-0.759	-0.784	CHTF8	
387	RNA815_10700	0.000	0.006	0.000	0.007	0.007	-6.087	-6.835	ENDOU	
388	RNA815_11580.1	0.002	0.006	0.001	0.007	0.007	-1.413	-2.250	PCNA	Neoblasts, gonads
389	RNA815_4903	0.004	0.007	0.004	0.007	0.007	-0.657	-0.613	HNRPLL	ovaries
390	RNA815_2969	0.000	0.007	0.000	0.007	0.007	-4.801	-6.816	PANX2	
391	RNA815_6011	0.004	0.006	0.004	0.007	0.007	-0.535	-0.750	ADRM1	
392	RNA815_589.1	0.003	0.005	0.003	0.007	0.007	-0.706	-1.267	LRRC23	
393	RNA815_6462.1	0.001	0.008	0.001	0.007	0.008	-2.985	-2.570	WSCD1	
394	RNA815_4082	0.003	0.005	0.004	0.007	0.007	-0.550	-0.837	IMPDH1	
395	RNA815_3333	0.004	0.006	0.004	0.007	0.007	-0.726	-0.556	PDIA4	
396	RNA815_23193	0.002	0.006	0.005	0.007	0.007	-1.370	-0.505	C11orf10	
397	RNA815_4580.1	0.005	0.008	0.004	0.007	0.008	-0.624	-0.797	GBAS	
398	RNA815_10420	0.000	0.008	0.000	0.007	0.008	-6.157	-5.031	PRKAR2A	
399	RNA815_3245	0.003	0.006	0.003	0.007	0.007	-0.888	-0.946	TSR1	
400	RNA815_847	0.000	0.012	0.000	0.006	0.012	-5.611	-4.464	PC	
401	RNA815_4471	0.004	0.007	0.002	0.006	0.007	-0.913	-1.458	KARS	
402	RNA815_1744	0.001	0.007	0.000	0.006	0.007	-3.240	-4.320	PYGB	
403	RNA815_3710.1	0.000	0.006	0.001	0.006	0.006	-4.248	-3.619	SLC6A7	
404	RNA815_293	0.000	0.009	0.000	0.006	0.009	-5.333	-4.771	ABCA5	
405	RNA815_4852.1	0.001	0.006	0.001	0.006	0.006	-2.423	-2.567	RRM2	
406	RNA815_11640	0.004	0.005	0.004	0.006	0.006	-0.605	-0.547	ALG13	
407	RNA815_3698	0.004	0.005	0.002	0.006	0.006	-0.595	-1.433	SLC25A12	
408	RNA815_777	0.004	0.006	0.003	0.006	0.006	-0.511	-1.114	ATP13A1	
409	RNA815_12817.1	0.005	0.009	0.004	0.006	0.009	-0.856	-0.501	PSMB2	
410	RNA815_1979.1	0.003	0.006	0.004	0.006	0.006	-0.904	-0.728	PDE1A	

Supplementary Table 1. Selected transcripts affected by irradiation and results of *in situ* hybridization. (*Continued*)

Rank	transcript	210Gi	CTRL1i	210G	CTRL1	Max	LOG2 Illumina	LOG2 Solid	Human Gene	ISH pattern
411	RNA815_1628	0.004	0.007	0.003	0.006	0.007	-0.565	-0.947	NAA15	
412	RNA815_2177	0.001	0.006	0.000	0.006	0.006	-3.154	-3.933	TTLL3	
413	RNA815_1752.1	0.004	0.006	0.004	0.006	0.006	-0.633	-0.559	LLGL1	
414	RNA815_15367	0.003	0.006	0.003	0.006	0.006	-0.848	-1.080	VBP1	
415	RNA815_1406	0.003	0.005	0.004	0.006	0.006	-0.621	-0.553	RARS	
416	RNA815_31446	0.005	0.009	0.004	0.006	0.009	-0.758	-0.613	RPS23	
417	RNA815_6245.1	0.000	0.008	0.000	0.006	0.008	-5.446	-4.469	MEGF9	
418	RNA815_2958	0.003	0.007	0.002	0.006	0.007	-1.270	-1.559	XRN2	
419	RNA815_1577.1	0.004	0.006	0.003	0.006	0.006	-0.575	-1.032	PKD2L1	
420	RNA815_2351	0.002	0.010	0.001	0.006	0.010	-2.642	-3.542	SND1	
421	RNA815_1566.1	0.004	0.007	0.003	0.006	0.007	-0.987	-1.049	ASAH2	
422	RNA815_4931	0.000	0.006	0.000	0.006	0.006	-5.984	-5.307	PLS1	
423	RNA815_2643	0.001	0.008	0.001	0.006	0.008	-2.538	-2.413	MCM4	
424	RNA815_34466	0.000	0.007	0.000	0.006	0.007	-7.770	-5.622	WSCD1	
425	RNA815_5181	0.003	0.006	0.003	0.006	0.006	-0.762	-1.224	EIF2B1	
426	RNA815_2673	0.000	0.005	0.000	0.006	0.006	-5.280	-5.295	KLHL18	
427	RNA815_5426	0.000	0.007	0.000	0.006	0.007	-5.496	-5.487	SRRM2	
428	RNA815_4930.1	0.004	0.007	0.003	0.006	0.007	-0.637	-0.913	NAP1L4	
429	RNA815_5405	0.003	0.024	0.001	0.006	0.024	-2.956	-2.843	ITGA2	
430	RNA815_6840.1	0.000	0.005	0.000	0.006	0.006	-3.534	-4.833	ENTPD6	
431	RNA815_199.1	0.004	0.011	0.002	0.006	0.011	-1.321	-1.295	ANKRD50	
432	RNA815_4354	0.004	0.007	0.003	0.006	0.007	-0.588	-1.085	ALPL	
433	RNA815_10065	0.000	0.006	0.000	0.006	0.006	-6.450	-7.539	PPP1CB	
434	RNA815_5196	0.004	0.005	0.004	0.006	0.006	-0.639	-0.531	EIF4A3	no pattern
435	RNA815_1583.1	0.001	0.009	0.000	0.006	0.009	-3.868	-5.207	WASL	
436	RNA815_170	0.004	0.008	0.001	0.006	0.008	-1.047	-2.010	ZNFX1	
437	RNA815_5747	0.004	0.007	0.004	0.006	0.007	-0.600	-0.625	DDX19A	
438	RNA815_34821	0.004	0.007	0.004	0.006	0.007	-0.779	-0.607	BOLA1	
439	RNA815_2262.1	0.003	0.006	0.003	0.006	0.006	-1.089	-0.974	RBBP4	
440	RNA815_3991	0.003	0.005	0.003	0.005	0.005	-0.568	-1.049	RHBDL2	
441	RNA815_1579	0.002	0.007	0.002	0.005	0.007	-1.648	-1.780	PWP2	
442	RNA815_3050	0.002	0.008	0.001	0.005	0.008	-1.726	-2.311	TGS1	
443	RNA815_1547.1	0.003	0.005	0.003	0.005	0.005	-0.683	-0.863	PAPOLG	
444	RNA815_15254.1	0.001	0.007	0.001	0.005	0.007	-2.913	-2.841	LITAF	
445	RNA815_472.1	0.004	0.008	0.002	0.005	0.008	-1.095	-1.455	EMR1	
446	RNA815_3729	0.002	0.012	0.001	0.005	0.012	-2.376	-2.119	NASP	
447	RNA815_1639	0.002	0.008	0.002	0.005	0.008	-1.714	-1.163	CSNK1E	
448	RNA815_107.1	0.001	0.007	0.001	0.005	0.007	-3.489	-2.942	DYSF	
449	RNA815_9476	0.000	0.007	0.000	0.005	0.007	-5.833	-6.401	PSMC5	
450	RNA815_5351	0.001	0.006	0.001	0.005	0.006	-2.103	-3.313	PHTF1	
451	RNA815_862	0.005	0.010	0.002	0.005	0.010	-1.086	-1.318	SUPT16H	gonads
452	RNA815_3916	0.000	0.007	0.000	0.005	0.007	-6.684	-3.683	FAM108A1	
453	RNA815_5089	0.002	0.005	0.001	0.005	0.005	-1.645	-2.254	BACE1	
454	RNA815_3042.1	0.001	0.005	0.001	0.005	0.005	-3.153	-2.461	MATN2	

210Gi, CTRL1i, 210G and CTRL1 – normalized expression in the corresponding library; Max – maximum expression; Log2 Illumina and Log2 SOLiD – Log2 change between expression level in control worms and irradiated worms according to the results from corresponding platform; Human Gene – human homologue of the transcript; ISH pattern – expression pattern, according to ISH.

Supplementary Table 2. Neoblast candidate gene list based on decline after irradiation and no changes between juveniles and adults.

Rank	transcript	IRR_vs_CTRL	Juv_vs_CTRL2	ISH	RNAi	HumanGene
1	RNA815_30797	-6.91	-	testis		UBE2G1
2	RNA815_11284.1	-6.41	-	testis		ADAM18
3	RNA815_27656.1	-6.21	-	epithelium		PDIA2
4	RNA815_3838.3	-6.13	-			PLS1
5	RNA815_5053.1	-6.12	-	gonads		GLIPR2
6	RNA815_20011.2	-6.01	-	testis		KREMEN1
7	RNA815_5453	-5.84	-	gonads	phenotype	KLHL10
8	RNA815_10430	-5.82	-	no pattern		ABCA9
9	RNA815_10705	-5.81	-	testis		DDO
10	RNA815_23405	-5.72	-	gonads	no fast phenotype	RPGR
11	RNA815_3584.1	-5.67	-	testis		VRK2
12	RNA815_5148	-5.66	-	gonads		TTLL10
13	RNA815_4577.4	-5.62	-	gut and gonads		SCNN1G
14	RNA815_16202	-5.61	-			CLPX
15	RNA815_233	-5.59	-	gut		ABCA8
16	RNA815_16357.1	-5.51	-	no pattern		PLSCR2
17	RNA815_1505.4	-5.5	-	gonads		AMAC1L3
18	RNA815_3473.1	-5.47	-	gonads		GPD2
19	RNA815_18995.1	-5.46	-	testis		CHD9
20	RNA815_7259.2	-5.46	-	testis		FAM78A
21	RNA815_25463	-5.39	-	gonads	no fast phenotype	ABCA5
22	RNA815_33579	-5.39	-			DDC
23	RNA815_30807	-5.37	-	testis		WSCD2
24	RNA815_4212	-5.35	-	testis		KIAA1324
25	RNA815_20881	-5.34	-	gonads and neoblasts	no fast phenotype	KLHL15
26	RNA815_12961	-5.18	-	gonads and neoblasts		RAB6A
27	RNA815_38928	-5.17	-	gonads		FHL3
28	RNA815_21261	-5.15	-	gonads		RPH3AL
29	RNA815_16738	-5.14	-	gut		GLIPR1L1
30	RNA815_3366.3	-5.11	-	testis		PLS1
31	RNA815_2673	-5.07	-	gonads		KLHL18
32	RNA815_6038	-5.05	-	testis		ANKRD44
33	RNA815_8923.1	-5.04	-	testis		NYNRIN
34	RNA815_13993	-5.03	-	gonads		SLC46A1
35	RNA815_25788	-5	-	gonads		AK1
36	RNA815_20234	-5	4.28	gonads and neoblasts		SYN1
37	RNA815_11841	-4.99	-	gut and gonads		RAB15
38	RNA815_14008	-4.99	-			GPD1
39	RNA815_7478.1	-4.95	-	gonads		PPIF
40	RNA815_10401	-4.9	-	gut and gonads		GLS
41	RNA815_9951	-4.83	-	gonads and neoblasts	no fast phenotype	AC217779.2
42	RNA815_21543	-4.77	-	testis		KREMEN1
43	RNA815_8998	-4.76	-	gonads	no fast phenotype	SRR
44	RNA815_12337.1	-4.76	-	ovaries		ZFP36L2
45	RNA815_25286	-4.75	-	gonads	no fast phenotype	CDK5RAP2
46	RNA815_36034	-4.74	-	gonads	no fast phenotype	3-Mar
47	RNA815_20399	-4.72	-	gonads		PHRF1
48	RNA815_1574.2	-4.71	-	testis		TSPAN18
49	RNA815_1534	-4.69	-	gonads and neoblasts	no fast phenotype	TTLL1
50	RNA815_4430.2	-4.69	-	testis		CSNK1E
51	RNA815_7955.2	-4.67	-	gonads and neoblasts	no fast phenotype	ZFP36L1
52	RNA815_13973	-4.67	-	testis		ATP11B
53	RNA815_8297	-4.65	-	gonads		SLC47A2
54	RNA815_12957	-4.65	-	gonads		APC2
55	RNA815_6970	-4.62	-	gut		GLIPR2
56	RNA815_6912	-4.61	-			PI16
57	RNA815_4558	-4.6	-	gonads		ACTB
58	RNA815_4783.1	-4.6	-	gonads		PPP1R3B
59	RNA815_12778	-4.58	-	gonads and neoblasts		PPPDE1
60	RNA815_2450.1	-4.57	-	gonads and neoblasts	no fast phenotype	BOLL

Supplementary Table 2. Neoblast candidate gene list based on decline after irradiation and no changes between juveniles and adults. (*Continued*)

Rank	transcript	IRR_vs_CTRL	Juv_vs_CTRL2	ISH	RNAi	HumanGene
61	RNA815_11308	-4.57	-	testis		CCNB1
62	RNA815_24349	-4.54	2.64	gonads and neoblasts		CAMKV
63	RNA815_18247.1	-4.54	-	no pattern		THAP6
64	RNA815_6266	-4.53	-	gonads	no fast phenotype	RQCD1
65	RNA815_369.3	-4.52	3.18	gonads		UNC13A
66	RNA815_9521.1	-4.5	-	gonads		CPSF4L
67	RNA815_15645	-4.5	-	gonads and neoblasts		SCRT2
68	RNA815_17489	-4.5	-			MCM7
69	RNA815_11462.1	-4.47	-	gonads		SLC18A1
70	RNA815_21306	-4.46	-	gonads		PRDX1
71	RNA815_13236.1	-4.46	-	gonads and neoblasts		CA12
72	RNA815_30114	-4.45	-	gonads and neoblasts	no fast phenotype	TTLL10
73	RNA815_30377	-4.41	-			CABP7
74	RNA815_10409.1	-4.41	-			KIFC1
75	RNA815_14904	-4.38	-	gonads	no fast phenotype	DNAJB8
76	RNA815_11498.1	-4.38	-			PLD1
77	RNA815_19389.1	-4.37	-			SLC16A2
78	RNA815_5305.1	-4.35	-	gonads	no fast phenotype	ZCWPW1
79	RNA815_17297	-4.29	-	gonads and neoblasts		AP2A1
80	RNA815_13879	-4.29	-	testis		ASCL1
81	RNA815_11327.1	-4.24	-	gonads		PPP1CA
82	RNA815_17987	-4.24	-	gonads		THBS1
83	RNA815_17618	-4.23	-	gonads		LPXN
84	RNA815_4673	-4.23	-	gonads and neoblasts		CD109
85	RNA815_12602	-4.21	-			CIZ1
86	RNA815_9198	-4.2	-	gonads	no fast phenotype	TMEM72
87	RNA815_14066	-4.19	-	testis		MYOF
88	RNA815_3208	-4.18	-	no pattern		TBR1
89	RNA815_14563	-4.17	-	ovaries		F11
90	RNA815_7816	-4.12	-	gonads		IDH3G
91	RNA815_19729.2	-4.12	2.69	gonads and neoblasts		CABP7
92	RNA815_18275	-4.12	-	gonads and neoblasts		MUC5B
93	RNA815_27827	-4.12	-	gonads and neoblasts		ZC3H4
94	RNA815_1597.3	-4.11	-	gonads and neoblasts		GLIPR2
95	RNA815_26608	-4.1	-	gonads		PPP1CC
96	RNA815_3837	-4.1	-	testis		TRPM4
97	RNA815_5787.1	-4.07	-	testis		NEK2
98	RNA815_6036	-4.03	-	testis		CBS
99	RNA815_6652	-4.03	-	testis		TUBA1B
100	RNA815_12596	-3.98	3.16	gonads and neoblasts		WBSCR27
101	RNA815_8696	-3.91	-	gonads and neoblasts	no fast phenotype	KIFC3
102	RNA815_30732	-3.91	-	testis		WSCD2
103	RNA815_28505	-3.9	-	gonads and neoblasts		GLIPR2
104	RNA815_5707	-3.88	-			KREMEN2
105	RNA815_4626	-3.86	-	gonads		CCNB1
106	RNA815_15326	-3.85	-	gonads	no fast phenotype	CCND2
107	RNA815_6973.1	-3.84	-	gonads	no fast phenotype	CTSL1
108	RNA815_13666.1	-3.84	-	testis		TUBA8
109	RNA815_23578	-3.83	-	gonads and neoblasts		GLIPR2
110	RNA815_2500.1	-3.81	-	gonads and neoblasts		ZFP36L2
111	RNA815_5834	-3.8	-	testis		SCN4A
112	RNA815_9892	-3.79	-	gonads and neoblasts	no fast phenotype	ZFP36
113	RNA815_9315	-3.79	-	gonads and neoblasts	no fast phenotype	AC010614.2
114	RNA815_2640	-3.76	-	gonads and neoblasts	phenotype	CPEB1
115	RNA815_11387.1	-3.75	-	testis		H1F0
116	RNA815_11522.1	-3.73	-			CEL
117	RNA815_17103.1	-3.71	-	gonads		HSPA6
118	RNA815_14005.1	-3.71	-	no pattern		BAG1
119	RNA815_6677	-3.68	-	no pattern		C1orf112
120	RNA815_5840	-3.66	-	gonads		MED4

Supplementary Table 2. Neoblast candidate gene list based on decline after irradiation and no changes between juveniles and adults. (*Continued*)

Rank	transcript	IRR_vs_CTRL	Juv_vs_CTRL2	ISH	RNAi	HumanGene
121	RNA815_19871.1	-3.66	-			RAX
122	RNA815_1323	-3.64	-	gonads	phenotype	RRM1
123	RNA815_4116.3	-3.61	-	testis		GLUL
124	RNA815_8566	-3.6	-	gonads and neoblasts	no fast phenotype	FLT1
125	RNA815_994.1	-3.6	-	no pattern		TRPA1
126	RNA815_27244	-3.57	-			SPAG8
127	RNA815_2897	-3.55	-	gonads	no fast phenotype	KPNA2
128	RNA815_2073	-3.5	-	ovaries		SLC26A5
129	RNA815_18983	-3.5	-	testis		ADAM18
130	RNA815_14143	-3.47	-	gonads		MEX3B
131	RNA815_59	-3.47	-	gonads		KRTAP10-7
132	RNA815_2177	-3.42	-			TTLL3
133	RNA815_23320.1	-3.41	4.02			H1F0
134	RNA815_13520.1	-3.4	-	gonads and neoblasts	no fast phenotype	TSPAN18
135	RNA815_3088.1	-3.36	-	gonads	no fast phenotype	CDC25A
136	RNA815_21514	-3.36	-	gonads and neoblasts		NAP1L1
137	RNA815_3365	-3.35	-	gonads and neoblasts		MAD2L1
138	RNA815_10105	-3.32	-	gonads and neoblasts		TMEM55A
139	RNA815_10706.1	-3.31	-	testis		KREMEN1
140	RNA815_14439.3	-3.28	-	testis		WSCD2
141	RNA815_813.1	-3.27	-	gonads and neoblasts		PIWIL1
142	RNA815_11198	-3.26	-	testis		CCNB1
143	RNA815_5862	-3.25	-	gonads	no fast phenotype	PCIF1
144	RNA815_28035	-3.19	-	gonads		SLC18A2
145	RNA815_2052.1	-3.09	-	gut and gonads		CAPN3
146	RNA815_9606	-3.07	-	testis		IDH2
147	RNA815_5941	-3.04	-	gonads	phenotype	CDC20
148	RNA815_6460	-3.03	-	gonads and neoblasts	no fast phenotype	TRIM3
149	RNA815_7369	-3.02	-	gonads		AC109322.1
150	RNA815_6391	-3.01	-			CDC42BPG
151	RNA815_20777.1	-3	3.82	gonads and neoblasts		SCARF1
152	RNA815_9816	-3	-			NTAN1
153	RNA815_4845	-2.99	-	gonads		SLC18A3
154	RNA815_6853	-2.99	2.46			SAFB2
155	RNA815_6217	-2.98	2.43	no pattern		RAD1
156	RNA815_12722	-2.97	-	gut and gonads		CAV3
157	RNA815_3335	-2.97	-	testis		NFXL1
158	RNA815_3206.1	-2.91	-	gut and gonads		ABCF2
159	RNA815_19501	-2.89	-	testis		CYCS
160	RNA815_16633	-2.88	-	gonads	no fast phenotype	AKAP1
161	RNA815_7206	-2.86	-	gonads	no fast phenotype	CDC25C
162	RNA815_5342	-2.85	-	gonads		MEX3B
163	RNA815_5405	-2.85	-	gonads	no fast phenotype	ITGA2
164	RNA815_108	-2.83	-	testis		COL7A1
165	RNA815_17643	-2.79	-			DSPP
166	RNA815_7117	-2.75	-	ovaries		PTMA
167	RNA815_31340	-2.72	-	testis		RBM12B
168	RNA815_15759	-2.7	-	gonads and neoblasts		PIWIL1
169	RNA815_4852.2	-2.7	2.69	gut and gonads		RRM2
170	RNA815_10824	-2.67	-	testis		CKS2
171	RNA815_8293	-2.54	-			PDP1
172	RNA815_5712.1	-2.52	2.01	gonads		CCNB1
173	RNA815_2351	-2.44	-	no pattern		SND1

IRR_vs_CTRL - Log2-change in expression after irradiation. Juv_vs_CTRL2 - Log2-difference in expression in juveniles, compared to adults; Values are only provided for the genes that are expressed significantly higher (4-fold difference or more) in juveniles, than in adults. For all the other genes in the list Log2-difference is between -2 and 2.

Supplementary Table 3. Neoblast candidate gene selection based on consistent decline during irradiation timecourse.

Rank	transcript	Max Percent	LOG2 D1/ CTRL2	LOG2 D3/ CTRL2	LOG2 D7/ CTRL2	ISH	RNAi	Probe number	Human Gene
1	RNA815_2025.1	0.379	-0.296	-0.588	-1.178	gonads	phenotype	Kay_2025	HSP90AA1
2	RNA815_1155.1	0.181	-0.482	-0.943	-2.006	gonads	no fast phenotype	B_36	DDX6
3	RNA815_15958	0.079	-0.243	-0.644	-0.686				HSPA8
4	RNA815_2657	0.061	-0.487	-1.493	-2.194	gonads	phenotype	B_14	HSPD1
5	RNA815_12477.2	0.055	-0.075	-0.441	-0.471				RPL11
6	RNA815_12035.1	0.055	-0.160	-0.306	-1.167				RPS17
7	RNA815_6753	0.043	-0.301	-0.569	-1.150	gonads and neoblasts	strong phenotype	B_51	DDX39B
8	RNA815_8346	0.041	-0.122	-0.506	-0.915				ALKBH3
9	RNA815_4101.1	0.038	-0.392	-0.818	-1.493	gonads		B_27	CCT5
10	RNA815_5606.1	0.028	-0.197	-0.938	-1.187	gonads		B_11	PURA
11	RNA815_7554	0.026	-0.208	-1.055	-1.542	gonads and neoblasts	no fast phenotype	7	PRMT1
12	RNA815_11859.1	0.025	-0.249	-0.394	-0.664	ovaries	no fast phenotype		RPL15
13	RNA815_2882.1	0.025	-0.421	-1.278	-1.821				CCT7
14	RNA815_21872	0.023	-0.348	-0.549	-0.635				RPS9
15	RNA815_1388	0.023	-0.325	-0.424	-1.098				EIF3C
16	RNA815_26208.2	0.023	-0.202	-0.238	-0.383				RPS13
17	RNA815_9146.1	0.022	-0.010	-0.755	-0.771				PSMB4
18	RNA815_12419	0.022	-0.387	-1.030	-1.264				PSMD8
19	RNA815_1672.2	0.021	-0.360	-0.547	-1.195				VCP
20	RNA815_1945.1	0.021	-0.325	-0.554	-1.392				CCT2
21	RNA815_3930	0.020	-0.696	-1.322	-1.656				TCP1
22	RNA815_5605.5	0.020	-0.846	-1.094	-3.642	gonads		24	TNRC18
23	RNA815_6491	0.020	-0.302	-0.931	-1.181				PSMD6
24	RNA815_813.1	0.019	-0.962	-1.184	-3.983	gonads and neoblasts		112	PIWIL1
25	RNA815_26500.1	0.019	-0.148	-1.008	-1.431				HSPE1
26	RNA815_3691.2	0.019	-0.160	-0.618	-0.762				PDIA6
27	RNA815_3268.1	0.018	-0.207	-0.815	-1.747	no pattern			CCT4
28	RNA815_935	0.017	-0.190	-0.595	-1.060				HSPA4L
29	RNA815_1350.1	0.016	-0.308	-0.914	-1.183				GALNT2
30	RNA815_1798.1	0.016	-0.092	-0.492	-1.111				RPN1
31	RNA815_3344	0.015	-0.201	-0.933	-1.043				STT3A
32	RNA815_7725.1	0.015	-0.865	-1.020	-3.660	gonads	no fast phenotype	22	
33	RNA815_1184	0.015	-2.167	-2.588	-5.735				IDE
34	RNA815_5606.2	0.015	-0.249	-1.047	-1.755	gonads		B_09	PURA
35	RNA815_3741.1	0.014	-0.577	-1.247	-2.241	gonads		B_41	CCT3
36	RNA815_8768	0.014	-0.267	-0.456	-0.926				RSL1D1
37	RNA815_1690	0.013	-0.621	-0.886	-1.386				PSMD2
38	RNA815_28505	0.013	-1.131	-1.552	-3.611	gonads and neoblasts		68	GLIPR2
39	RNA815_10026	0.013	-0.466	-0.757	-1.119				PSMD14
40	RNA815_11581	0.013	-0.377	-0.692	-1.243	gonads	no fast phenotype	B_43	CCT8
41	RNA815_5034	0.013	-0.522	-1.176	-1.748	gonads		B_24	FKBP4
42	RNA815_310.1	0.012	-0.620	-1.064	-2.156				TDRD9
43	RNA815_2969	0.012	-0.916	-1.444	-3.961				PANX2
44	RNA815_2897	0.012	-1.042	-1.373	-4.218	gonads	no fast phenotype	48	KPNA2
45	RNA815_10718	0.012	-1.912	-2.394	-5.426	testis		16	FAM78B
46	RNA815_9864.1	0.012	-1.588	-1.994	-8.437	gonads		11	FAM78B
47	RNA815_4445	0.012	-0.373	-0.773	-1.698				NCL
48	RNA815_5205	0.011	-0.877	-1.837	-1.885	no pattern	no fast phenotype	Kay_5205	DKC1
49	RNA815_13247	0.011	-0.252	-0.473	-0.616				RP11-111K18.1

Supplementary Table 3. Neoblast candidate gene selection based on consistent decline during irradiation timecourse. (*Continued*)

Rank	transcript	Max Percent	LOG2 D1/ CTRL2	LOG2 D3/ CTRL2	LOG2 D7/ CTRL2	ISH	RNAi	Probe number	Human Gene
50	RNA815_1386.1	0.011	-0.283	-0.890	-1.843				NCL
51	RNA815_1734	0.011	-1.478	-1.751	-3.041				ARIH1
52	RNA815_4172.1	0.010	-0.224	-0.769	-1.384				TARDBP
53	RNA815_7841	0.010	-0.154	-0.250	-0.317				GSN
54	RNA815_1521	0.010	-1.605	-2.385	-6.262				DNAJB7
55	RNA815_325	0.010	-0.060	-0.336	-0.890				GANAB
56	RNA815_2207.1	0.010	-0.268	-0.742	-0.877				DPP3
57	RNA815_4315.1	0.009	-2.275	-2.456	-6.182				PPEF1
58	RNA815_8006	0.009	-1.726	-2.120	-5.665				CAPN11
59	RNA815_15286	0.009	-0.236	-1.124	-1.367	gonads	no fast phenotype	B_19	CCT8
60	RNA815_5162.1	0.009	-0.633	-1.279	-1.827				NOP56
61	RNA815_5712.1	0.009	-0.976	-1.488	-2.469	gonads		60	CCNB1
62	RNA815_19501	0.009	-0.972	-1.979	-3.623	testis		72	CYCS
63	RNA815_11799	0.009	-0.057	-0.251	-1.786				SSB
64	RNA815_31446	0.009	-0.032	-0.158	-0.213	ovaries	no fast phenotype		RPS23
65	RNA815_6973.1	0.009	-1.332	-1.559	-3.951	gonads	no fast phenotype	109	CTSL1
66	RNA815_6256.1	0.009	-0.174	-1.029	-1.307				PSMC4
67	RNA815_5016.2	0.009	-0.833	-1.965	-2.024				EXOSC10
68	RNA815_6011	0.008	-0.314	-0.824	-0.952				ADRM1
69	RNA815_9365	0.008	-1.009	-1.705	-2.984				UBE2R2
70	RNA815_2074	0.008	-0.691	-0.983	-2.132				MYO1C
71	RNA815_2344.1	0.008	-0.408	-0.643	-1.043				USP14
72	RNA815_11580.1	0.008	-0.945	-1.981	-2.145	gonads and neoblasts	phenotype	10	PCNA
73	RNA815_1744	0.008	-1.347	-1.376	-3.631				PYGB
74	RNA815_976	0.008	-0.017	-0.246	-1.813				NPC1
75	RNA815_409	0.008	-1.081	-1.113	-2.517	gonads		B_54	TOP2A
76	RNA815_27193	0.008	-0.022	-0.150	-0.447				PPIF
77	RNA815_7489.1	0.008	-1.025	-1.434	-2.634	testis	phenotype	5	CDK1
78	RNA815_2958	0.007	-0.734	-0.841	-2.040				XRN2
79	RNA815_28959	0.007	-0.343	-0.628	-1.204				NHP2L1
80	RNA815_3152	0.007	-0.104	-0.533	-1.122				NSF
81	RNA815_980	0.007	-0.526	-1.424	-3.757				PIWIL1
82	RNA815_7068	0.007	-1.030	-1.266	-1.553				GPN1
83	RNA815_8386.1	0.007	-0.797	-1.967	-2.983				NOLC1
84	RNA815_11700	0.007	-1.124	-1.792	-2.391				UBE2S
85	RNA815_14114	0.007	-0.147	-0.478	-0.954				MRPL47
86	RNA815_7009	0.007	-0.041	-1.256	-1.950				ADIPOR2
87	RNA815_10824	0.007	-1.138	-2.152	-2.613	testis		32	CKS2
88	RNA815_1704	0.007	-0.257	-0.276	-2.473	gonads and neoblasts		35	SEL1L
89	RNA815_9366	0.006	-0.614	-0.815	-1.071				BRIX1
90	RNA815_7479	0.006	-0.005	-0.542	-0.609				TXNL1
91	RNA815_4852.1	0.006	-0.483	-0.709	-2.488				RRM2
92	RNA815_10807	0.006	-1.147	-1.378	-6.044	testis		13	PRKAR2B
93	RNA815_10065	0.006	-1.653	-2.420	-7.668	testis		12	PPP1CB
94	RNA815_3317	0.006	-0.627	-1.293	-2.694				ATAD3B
95	RNA815_12848.2	0.006	-0.477	-1.434	-1.514				TBCB
96	RNA815_17292	0.006	-1.912	-2.901	-6.579				PI16
97	RNA815_16106.1	0.006	-0.848	-1.437	-2.478	gonads		34	CKS1B
98	RNA815_4362.1	0.006	-0.513	-0.685	-2.529				MYH8
99	RNA815_14564.1	0.006	-0.600	-0.913	-1.533				TRMT112
100	RNA815_2127	0.006	-0.147	-0.387	-1.129				CYP20A1
101	RNA815_4929.5	0.006	-0.163	-0.257	-0.268				ACTG1
102	RNA815_5227.1	0.006	-0.826	-0.955	-1.228				API5
103	RNA815_20103	0.006	-2.042	-2.436	-6.109				ITPA

Supplementary Table 3. Neoblast candidate gene selection based on consistent decline during irradiation timecourse. (*Continued*)

Rank	transcript	Max Percent	LOG2 D1/ CTRL2	LOG2 D3/ CTRL2	LOG2 D7/ CTRL2	ISH	RNAi	Probe number	Human Gene
104	RNA815_909	0.005	-0.441	-0.995	-2.380	gonads and neoblasts	no fast phenotype	38	MCM6
105	RNA815_2525	0.005	-0.031	-1.076	-1.209				SDAD1
106	RNA815_13120.2	0.005	-1.010	-1.388	-2.867				CSNK1G1
107	RNA815_3603	0.005	-0.476	-0.722	-1.769	gonads		2	PLK2
108	RNA815_2495	0.005	-0.373	-0.713	-3.298				TDRD1
109	RNA815_7946	0.005	-0.276	-0.520	-0.666				HSD17B10
110	RNA815_13797	0.005	-1.122	-1.344	-2.432				FBL
111	RNA815_443	0.005	-0.428	-0.449	-0.805				RNF20
112	RNA815_940	0.005	-0.026	-0.126	-1.000				UPF1
113	RNA815_1959	0.005	-0.376	-1.242	-1.870				NCBP1
114	RNA815_3705	0.005	-1.008	-1.305	-1.707	gonads	phenotype	3	BOP1
115	RNA815_2981.1	0.005	-0.082	-0.686	-1.211				EDEM2
116	RNA815_6247	0.005	-0.768	-1.091	-2.682				PPAP2A
117	RNA815_9739.2	0.005	-0.135	-0.334	-0.635				SUB1
118	RNA815_21836	0.005	-0.088	-0.331	-1.054				POLR2E
119	RNA815_5212	0.005	-2.178	-3.511	-4.248				CHML
120	RNA815_4996.1	0.004	-0.818	-1.375	-1.724				PDZRN4
121	RNA815_5941	0.004	-1.182	-1.338	-4.561	gonads	phenotype	17	CDC20
122	RNA815_5999	0.004	-1.648	-1.783	-5.133				DNAJB2
123	RNA815_3088.1	0.004	-1.226	-1.914	-2.996	gonads	no fast phenotype	27	CDC25A
124	RNA815_1412	0.004	-0.524	-2.331	-2.807	gonads		29	INCENP
125	RNA815_16818	0.004	-0.209	-0.547	-1.914				UBE2I
126	RNA815_7812	0.004	-0.157	-0.582	-1.112				BCKDHA
127	RNA815_3289	0.004	-1.816	-2.219	-3.588				HPS1
128	RNA815_10739.1	0.004	-0.366	-0.616	-0.792				ACOT8
129	RNA815_15482	0.004	-0.138	-0.525	-0.803				WDR43
130	RNA815_32932	0.004	-0.514	-0.939	-1.455				SNRPE
131	RNA815_2942	0.004	-1.032	-1.250	-2.391				AC010872.2
132	RNA815_1140.1	0.004	-0.377	-0.412	-1.082				KDM2A
133	RNA815_13120.1	0.004	-0.391	-0.954	-1.558				CSNK1G1
134	RNA815_9198	0.004	-0.819	-0.855	-3.684	gonads	no fast phenotype	21	TMEM72
135	RNA815_8039	0.004	-0.265	-0.576	-0.888				MRPS5
136	RNA815_6774	0.004	-0.605	-0.626	-0.848	gonads	no fast phenotype	B_26	DDX52
137	RNA815_3220.1	0.004	-0.130	-0.598	-1.183				ADCK4
138	RNA815_5862	0.004	-1.095	-1.622	-3.657	gonads	no fast phenotype	23	PCIF1
139	RNA815_4084.1	0.004	-0.167	-0.196	-0.580				METTL13
140	RNA815_7120	0.004	-0.555	-0.590	-1.507				LANCL2
141	RNA815_2461	0.004	-0.140	-0.222	-1.895				THOC5
142	RNA815_10450	0.004	-1.031	-1.094	-1.663				SH3BP5L
143	RNA815_6053	0.004	-0.714	-0.989	-2.866	testis		28	PAPD4
144	RNA815_6266	0.004	-2.322	-2.421	-4.381	gonads	no fast phenotype	19	RQCD1
145	RNA815_19139	0.004	-1.022	-1.818	-2.103	gonads		B_32	H2AFV
146	RNA815_9275	0.004	-0.064	-0.863	-2.427	gonads and neoblasts		37	CEP128
147	RNA815_5919	0.004	-0.322	-0.925	-1.287				SENP7
148	RNA815_15758	0.004	-0.248	-1.419	-2.368				NAA40
149	RNA815_24740	0.004	-0.447	-0.691	-0.885				LSM5
150	RNA815_26601	0.004	-1.326	-2.005	-2.803	gonads		30	RPA3
151	RNA815_10490	0.004	-1.453	-2.803	-5.476				UFD1L
152	RNA815_8952	0.004	-0.285	-0.450	-1.422				PAK1IP1
153	RNA815_7206	0.003	-0.613	-0.732	-2.364	gonads	no fast phenotype	127	CDC25C
154	RNA815_5434	0.003	-0.113	-0.331	-0.813				ACPL2

Supplementary Table 3. Neoblast candidate gene selection based on consistent decline during irradiation timecourse. (*Continued*)

Rank	transcript	Max Percent	LOG2 D1/ CTRL2	LOG2 D3/ CTRL2	LOG2 D7/ CTRL2	ISH	RNAi	Probe number	Human Gene
155	RNA815_12985	0.003	-0.279	-0.428	-0.864				CIAO1
156	RNA815_3368	0.003	-1.546	-2.105	-2.830				ABCG2
157	RNA815_9236	0.003	-0.771	-0.996	-1.100				COX11
158	RNA815_27599	0.003	-0.574	-1.047	-1.619				FAM136A
159	RNA815_6146.1	0.003	-0.604	-0.792	-2.503	gonads		33	
160	RNA815_2171	0.003	-0.777	-1.329	-2.233				TMEM209
161	RNA815_12614	0.003	-0.315	-0.734	-0.749				HPDL
162	RNA815_14115	0.003	-0.961	-1.681	-1.877				SURF2
163	RNA815_9420	0.003	-0.766	-1.146	-1.227				DNAJC8
164	RNA815_10300	0.003	-0.502	-1.023	-1.157				UNG
165	RNA815_11459	0.003	-0.583	-1.224	-2.457	testis		36	ZNHIT6
166	RNA815_16633	0.003	-1.810	-3.037	-3.478	gonads	no fast phenotype	25	AKAP1
167	RNA815_38759	0.003	-0.161	-0.954	-1.000				TIMM8A
168	RNA815_9071	0.003	-0.546	-1.299	-1.967				TRMT6
169	RNA815_14320.1	0.003	-0.776	-0.825	-1.393				NAA20
170	RNA815_17953.1	0.003	-0.255	-0.523	-0.657	ovaries	no fast phenotype		MRPL14
171	RNA815_2058	0.003	-0.663	-1.322	-1.763				RRBP1
172	RNA815_1952.1	0.003	-0.127	-0.587	-1.048				ASCC2
173	RNA815_1456	0.003	-0.569	-0.939	-1.436				MYH9
174	RNA815_12820	0.003	-0.212	-1.029	-1.286				PABPN1
175	RNA815_1583.2	0.003	-1.934	-2.069	-5.897			15	TPRXL
176	RNA815_6587	0.003	-0.063	-0.562	-1.174	ovaries	no fast phenotype		PRMT3
177	RNA815_12305	0.003	-0.359	-1.281	-1.372				CBX1
178	RNA815_358.1	0.003	-0.585	-0.768	-1.655				PIF1
179	RNA815_2497	0.003	-0.938	-1.000	-3.650	gonads and neoblasts	no fast phenotype	1	MCM2
180	RNA815_3040	0.003	-0.822	-1.154	-3.307				MCMBP
181	RNA815_8772.2	0.003	-2.066	-4.268	-5.175				PRDX3
182	RNA815_14727	0.003	-0.442	-0.857	-1.778				SRSF7
183	RNA815_10438.1	0.003	-0.869	-0.989	-1.196				A4GNT
184	RNA815_8700	0.003	-0.960	-1.477	-2.395	gonads		39	RFC5
185	RNA815_3205	0.003	-0.055	-0.837	-1.268				CPSF3
186	RNA815_1710.1	0.003	-0.410	-0.805	-2.363				PRRC2C
187	RNA815_7061	0.003	-0.562	-0.825	-1.431				MRPS27
188	RNA815_1334	0.003	-0.767	-0.995	-1.995				TRPC4AP
189	RNA815_5794.1	0.003	-0.820	-0.848	-1.147				WDR45L
190	RNA815_12898	0.003	-1.279	-1.835	-5.923	gonads		14	CHMP4B
191	RNA815_5093	0.003	-0.663	-1.195	-1.756				BCKDK
192	RNA815_8897	0.003	-0.354	-0.966	-1.745				TOMM40
193	RNA815_22216	0.003	-1.527	-1.691	-2.896				MAP7D3
194	RNA815_4556	0.003	-0.085	-0.500	-1.021				WDR46
195	RNA815_9319.1	0.003	-0.266	-0.995	-1.374	gonads		B_44	PPIL1
196	RNA815_3377	0.003	-0.828	-1.706	-2.247				SND1
197	RNA815_10991	0.003	-0.419	-0.702	-1.719				DAP3
198	RNA815_8452	0.003	-0.228	-0.361	-1.625				MRPS22
199	RNA815_2693.1	0.003	-0.335	-1.365	-2.000	gonads	no fast phenotype	B_29	HMGB3
200	RNA815_17954.1	0.003	-0.677	-0.863	-1.905				CLPP
201	RNA815_19555	0.003	-0.237	-1.251	-1.376				GSTK1
202	RNA815_5534	0.003	-0.076	-0.087	-1.928				BMS1
203	RNA815_1474	0.003	-0.076	-0.219	-1.690				DHX57
204	RNA815_13849	0.003	-0.656	-1.222	-1.340				TK2
205	RNA815_5047	0.003	-0.863	-1.496	-2.190				C4orf29
206	RNA815_14775	0.003	-0.766	-1.082	-1.231				METTL7B
207	RNA815_14787	0.002	-0.697	-0.735	-1.084				UTP11L
208	RNA815_2355	0.002	-1.117	-1.463	-1.838				CDC16

Supplementary Table 3. Neoblast candidate gene selection based on consistent decline during irradiation timecourse. (*Continued*)

Rank	transcript	Max Percent	LOG2 D1/ CTRL2	LOG2 D3/ CTRL2	LOG2 D7/ CTRL2	ISH	RNAi	Probe number	Human Gene
209	RNA815_5962.1	0.002	-0.220	-0.266	-0.865				TSSC1
210	RNA815_9671	0.002	-1.158	-1.186	-2.688	no pattern		31	RAD51
211	RNA815_6092	0.002	-0.121	-0.314	-1.399				ILF2
212	RNA815_7998	0.002	-0.772	-1.122	-1.585				POLR2G
213	RNA815_5592	0.002	-0.570	-1.393	-1.634				HNRNPH2
214	RNA815_7602	0.002	-0.196	-0.454	-1.977				PRPF19
215	RNA815_7126.1	0.002	-0.037	-0.817	-0.917				NUS1
216	RNA815_15759	0.002	-0.521	-0.874	-3.307	gonads		111	PIWIL1
217	RNA815_14133	0.002	-0.698	-1.117	-1.443				FBXO17
218	RNA815_16476	0.002	-0.848	-1.037	-1.478				MRPL28
219	RNA815_7397	0.002	-0.472	-1.372	-1.560				OGG1
220	RNA815_6460	0.002	-1.132	-1.585	-3.833	gonads and neoblasts	no fast phenotype	20	TRIM3
221	RNA815_12753.1	0.002	-0.823	-1.845	-2.322				TPX2
222	RNA815_6772	0.002	-0.689	-1.269	-2.139				WDR61
223	RNA815_17053	0.002	-1.951	-2.195	-4.443	gonads		18	CABLES2
224	RNA815_18550.1	0.002	-0.174	-0.461	-1.737	no pattern	no fast phenotype	Kay_18550	TBP
225	RNA815_20641	0.002	-1.120	-1.820	-2.691				MPV17L2
226	RNA815_6967	0.002	-0.673	-2.083	-2.287				MRPL37
227	RNA815_15785	0.002	-0.355	-1.212	-1.290				GPN3
228	RNA815_3166	0.002	-0.671	-0.754	-2.178				NUP155
229	RNA815_14061	0.002	-1.329	-2.263	-2.585				FBXO43
230	RNA815_16079	0.002	-0.571	-1.100	-1.705				SLBP
231	RNA815_5773	0.002	-1.222	-1.618	-2.322				SMC2
232	RNA815_4217.1	0.002	-0.090	-0.511	-0.642				UTP18
233	RNA815_7361	0.002	-0.141	-0.169	-0.712				ACBD5
234	RNA815_16774	0.002	-0.654	-0.882	-1.841				MRPL46
235	RNA815_3687	0.002	-0.719	-0.753	-0.934				AC105345.1
236	RNA815_10393	0.002	-0.404	-0.989	-1.596				PTPLAD1
237	RNA815_6157.1	0.002	-0.420	-0.697	-1.592				SON
238	RNA815_21516	0.002	-0.149	-0.882	-1.900				DCP1B
239	RNA815_10919.1	0.002	-0.564	-1.697	-1.720				PUS1
240	RNA815_2792	0.002	-0.363	-0.544	-1.336				POLR3A
241	RNA815_5898	0.002	-0.443	-0.920	-1.785				DCAF17
242	RNA815_8716	0.002	-0.204	-0.763	-1.654				MRPL38
243	RNA815_9318.1	0.002	-0.384	-1.501	-3.062				NEFH
244	RNA815_2621	0.002	-0.200	-1.624	-3.209	gonads and neoblasts		26	TRIM59

Log2 D1/CTRL2, D3/CTRL2 and D7/CTRL2 – Log2 change in expression level between irradiated worms at first, third or 7th day after irradiation respectively and control worms. Human Gene – human homologue of the transcript; RNAi – stem cell related phenotypes caused by knockdown. ISH – expression pattern determined by ISH. Transcripts with no clear expression pattern are highlighted in grey; the ones expressed in differentiated tissues – in red; in testes – in green; in ovaries – in magenta; gonad-specific – in blue; genes expressed in neoblasts and gonads – in dark blue. Log2 D1/CTRL2, D3/CTRL2 and D7/CTRL2 – Log2 change in expression level between irradiated worms at first, third or 7th day after irradiation respectively and control worms.

Supplementary Table 4. Primers used for ISH probe templates amplification.

ID	Orient	Sequence 5'-->3'	Size
3206.1	FORW	ATTTCATGGGCTCAGACGAC	755
3206.1	REV	GCCGTTGAGAAAGTCCTGAG	
233	FORW	TGTTCAGGATGTCAGCTTCG	755
233	REV	CGATAACGCTGCTGTTTGAA	
4558	FORW	GCAACGCCTACAGAAAGAGC	766
4558	REV	TTGACAGCTTCATGCACACA	
7369	FORW	CGAAACCTACTGCGTCAAGA	751
7369	REV	AAGAGCTGGTGGAACGAAAA	
25788	FORW	AGCTGTCACATTTGGCTTCA	666
25788	REV	AATGCCAACGTCCTCTTTGT	
4577.4	FORW	GGCTGCATTAAAAACCGAAG	751
4577.4	REV	CAGGTTGCTGAAAAGCAGTG	
994.1	FORW	GCAATCTGCTTCAACGACAG	747
994.1	REV	CCTTCGGAAATCCTCCAAAT	
6038	FORW	CGGCCAGATACTGACACTGA	749
6038	REV	AGTTCCGAGCTGCTCTTTGA	
2897	FORW	ATCTCCTGCCGATCAGTTTG	750
2897	REV	TGGTCTGTGCCAGTCACAAT	
20881	FORW	ATCTGCTTGCAGGATTCGAG	757
20881	REV	AACAGCAGGCATTCGTGAGT	
2673	FORW	TGTTCGAAGTTGACCAGCAG	752
2673	REV	CCAAGCAATGAGCAAGTCAA	
14066	FORW	GTTGGACCCAGAGAGAGACG	750
14066	REV	GCTCAGCACTTTTCGTAGCC	
26608	FORW	GCTGCTTCCAGTGGCTACC	653
26608	REV	CCTAAGGCCAGCGTCAAGTA	
11327.1	FORW	AATCAACATTGTGGGCGACT	750
11327.1	REV	TGTTGTAGACGCCGCAGTAG	
2052.1	FORW	GGACTCGAGACACGCCTACT	749
2052.1	REV	GCTTTAGCTTGTGCCGAAAG	
12722	FORW	AGGTGTTCAGCTCCACGAAG	729
12722	REV	GGCTACTCCAGCCAATGAGA	
20399	FORW	GCCTTCTCGCTCGTAGACC	783
20399	REV	AGACGAAGCAGATCGATTCAC	
23405	FORW	CACCAATGGAGAATCCAACC	725
23405	REV	GGCGCAGTCCTTTAGGAAAC	
21306	FORW	TGTCGTCGAATTATCGCAAA	741
21306	REV	ATCCGACAGTCGAGTCCAAG	
11308	FORW	TGCAATCAGTGGTTTGGTTG	760
11308	REV	CACGAATCAACAGCCAGGTA	
5712.1	FORW	GGAGCCATAGCTGCAGAGTT	745
5712.1	REV	CTGGATGTAGGGCGGATAGA	
15326	FORW	GACCGCTATCTCAGCAAACC	766
15326	REV	CTCGGGAACTTCAGGTCAAA	
4626	FORW	TTAACGACTGGCTGTTGCAG	748
4626	REV	ACAGCGCAGCTGAATCAGTA	
11198	FORW	CAGCTGGCCTTCTACTCTGG	752
11198	REV	CTGGGTACAGTGGGGACAAT	
7478.1	FORW	GGGTGTCCACTTTCTGGATG	735
7478.1	REV	GTCCGAAACACAGGAGACCA	
5053.1	FORW	GCAGTCATTTTGCCCAAGTT	749
5053.1	REV	GTACTCGTTGTGCGAGACCA	
1597.3	FORW	TGTCAACATGTGGTACTATGAAG	401
1597.3	REV	TGGGCTAAAATCGTCGAAAT	
6912	FORW	TTGGCCCGGTATAAAAACAG	756
6912	REV	TTTTACCGGACCGACAATTC	
28505	FORW	CTTTGTGTTGGCAGTGCAAG	657
28505	REV	TGTAATGTAATTGCCTCCGTTG	
6970	FORW	ATGGAGTTGCTGTCGGTTTC	749
6970	REV	GATCCTTCTGGCAGCTCAAC	

Supplementary Table 4. Primers used for ISH probe templates amplification. (*Continued*)

ID	Orient	Sequence 5'-->3'	Size
16738	FORW	TAATCACGCCACTGCAAATC	761
16738	REV	ACGGGAACTCGTTGTTGAAG	
23578	FORW	CGCTCGGTGACGTTCAG	723
23578	REV	CATCATCATTGGCTCCTTGA	
19501	FORW	GCCAACAAGAGCAAGGGTAT	636
19501	REV	CCAAATTGATGCCCAGTTTC	
14904	FORW	TCACAACGACTGCTTGCTTC	752
14904	REV	GGCGTACTCGAGTGGTCTTC	
30377	FORW	TCTTGTTGATGCTGGAATGG	620
30377	REV	TGAGTCGACTAATTCGAAATGC	
19729.2	FORW	CGACTAATTCGAAATGCCAAC	542
19729.2	REV	GAGCGAGAGAAACAAATGAAAAA	
13236.1	FORW	TTCTGCACCAAGCATGACTC	754
13236.1	REV	CACTGTCGGTGCATCTTCAT	
3473.1	FORW	GACCCAAACAAGCAGGACAC	758
3473.1	REV	GCCTGTGATCCGAGTCCTTA	
10705	FORW	TGGCAGTTCAATTGCTTGAC	748
10705	REV	GCTCGAGACCAGATTTTGGA	
4212	FORW	TAACGATGGCGAAGGAAAAC	751
4212	REV	CGGCTGATTTTTCAGGTGAT	
10401	FORW	GCAACTCCTTCTTCCTGGAC	628
10401	REV	CGTTCTCGCGCATGTAGTAG	
4116.3	FORW	CCCGGTTAGTTGATCAAAGC	712
4116.3	REV	CCAGCAATTTTCACACCAGA	
13879	FORW	TCATGTCTGAGCAGCTTGGT	751
13879	REV	TCTTGCTGGCATTGAGTCAC	
19871.1	FORW	TTGGAGGCTAGGCAGAGAAA	766
19871.1	REV	GAAAGGCGCTAAAAACGATG	
3365	FORW	ACCAATGGTGGCAGAGAAAG	739
3365	REV	CCCATTCACAGGGGATAGAA	
8566	FORW	AGCTGTAGCCAGCGAGTTGT	748
8566	REV	GAGCTGATTCTGCTGCTCCT	
8923.1	FORW	AGAGATGGGCGATTTTCCTT	752
8923.1	REV	CGATTCAATCAAGCCGTTCT	
5834	FORW	TACCTTTGCGGGAAGTTGAC	742
5834	REV	CCAATTGTCCGTGTTGTCTG	
7816	FORW	GATTCTTGGCTGCCCTATCA	744
7816	REV	GGCAAACGTCTAGGAACAGG	
5453	FORW	CGTGGTCCGAGTTGACTACA	747
5453	REV	GCTCGCTCCTGGTAGTTGAA	
24349	FORW	AACCAGCGCCAGTATGAAGT	744
24349	REV	GGCGAAGTAAAATTAGAAGA	
5342	FORW	TATTCGTCCGCCGATTTAAC	746
5342	REV	AAAGTACGCGACCCAAACTG	
14143	FORW	ATTCGGCTTCGTAGTGCTGT	716
14143	REV	CTTGATCCGCTGACCATTTT	
10409.1	FORW	GAGGAACAGCCGGTTGAAG	748
10409.1	REV	CTTTCAATTTGACGGCGTTT	
8696	FORW	TAGCGGAGCATCTCCAACTT	754
8696	REV	TCCTCTTTCGCAATTGAAGC	
17618	FORW	GTCAGCCGGGGTCAATATC	761
17618	REV	GCCTCCTCCTTGGTCTTCTC	
38928	FORW	GAAGGTTAAACGCTGCAAGG	489
38928	REV	TCATGCAGACACCCACTAGC	
11387.1	FORW	GGCCAAGAGGACCAAGAAG	748
11387.1	REV	TGTTGCTTTGTCAGCTGCTC	
23320.1	FORW	CTCTCCTCCCTCCCCTTCT	711
23320.1	REV	GTGCCTTCCTGGTCCTCTT	
28035	FORW	AACAGCCGTCGAGAAATCAT	640
28035	REV	GCTTTCGCTCGTTCGACTT	

Supplementary Table 4. Primers used for ISH probe templates amplification. (*Continued*)

ID	Orient	Sequence 5'-->3'	Size
4845	FORW	TACAAGACCGCCTTCTGCTT	742
4845	REV	TGCAAATTAATGCACAGCAA	
13993	FORW	AGCGTACGGCATCCAAATAC	746
13993	REV	TCCCATGCTGGCATAGTACA	
11462.1	FORW	GAAGTCGAACGAGCAAAAGC	748
11462.1	REV	CTCCGTACTCGGCTGTGAA	
8297	FORW	TGCAGGCTGTATTCATCTGG	752
8297	REV	CTCCATGTCGGTCTCCATCT	
4673	FORW	AGATGCGCTAACCCTGTCAC	750
4673	REV	GTCAGGCCAGTTCTGTTGGT	
25286	FORW	GGACTGCAAGCAAGTCCATT	723
25286	REV	CGCCATTCCTCACTGTGTATT	
14008	FORW	CTGGTGCTGCTGATGGTTC	732
14008	REV	TACACCTCGGCAGCAGTCT	
21514	FORW	GTGGTGCAAAAGGCAAGAAT	730
21514	REV	GGTGCTGGGTAGAAGTTTGC	
14563	FORW	TTGCCTTTCGGAAGACTGAT	755
14563	REV	TTGAGCGGCTTGGTAAAACT	
6973.1	FORW	GACTGGGCGTCAATCAGTTT	751
6973.1	REV	TTGTCTCGGGCAATCCTAAC	
1505.4	FORW	CAGGCACAGCTGTCTGAAAT	743
1505.4	REV	CTGTTGGGTGTTGTTTGCAG	
15759	FORW	CTTTGCCAGAGGCAATGAG	753
15759	REV	GTCAACGCGCATGATCTTAT	
813.1	FORW	ATTTCGTTCGGATCAACACC	748
813.1	REV	ATCAGAATCGGCTGCTCACT	
12778	FORW	CAAGCCGAAAGACGTCAACT	747
12778	REV	TGGCGAAAAAGGGTTGATAA	
6391	FORW	GCCTGTCGAAGTCTTTGCTC	737
6391	REV	GCAGTAGGCTCAGCAGCAG	
4430.2	FORW	CTTCATGGTTGGGGATCAGT	742
4430.2	REV	CAGCTTCATAAGGCCGGTAA	
5787.1	FORW	CTTAGGAAGCGGAACGTTTG	745
5787.1	REV	GTGATCGAGGACCACCATTT	
8293	FORW	GTGCCAGGAGTTCTTGGAGA	756
8293	REV	CGACATTCGCTACAAGTGGA	
4783.1	FORW	CAAGCAGCAACAAACGAGAC	747
4783.1	REV	TTCAGATTGGTGCAGCAGAG	
5305.1	FORW	GCAAGGCTTCATCACTCGAT	749
5305.1	REV	CGGCTTTGATTTGGCATTAT	
5606.1	FORW	GTTCTATCTGGACGTGAAGC	587
5606.1	REV	ATGTACTCGTCCAGGATGTC	
2025.1	FORW	TGGCCTCTATAAAATCTTCG	603
2025.1	REV	TCCTCCTTCTCTGTGAGGTA	
5034	FORW	AAAGGCGTTATTAAGCACTG	602
5034	REV	GCATTCAGTTTAATGGCTTC	
19139	FORW	TCAAATTGGAGCTACTGGAT	600
19139	REV	CCGACCATGTAGAAATAGGA	
2693.1	FORW	TGGTGAAGGACAAGATGAGG	513
2693.1	REV	GCTTTGTTGGGACGAAGGT	
33579	FORW	ATTCCTCTTTTGCCGATGC	588
33579	REV	TGTCGTAGCCGACATTGAC	
8998	FORW	CATGTCAACCAGCTGTCACC	743
8998	REV	ACTCGTCGGACGTATTCAGC	
6036	FORW	CTGCGTCACTTGAACTGCAT	755
6036	REV	GCAGACCTTTAATCGGACCA	
12961	FORW	GCGGTAACAGAGGCATCAAT	747
12961	REV	GCACACTGAGCTGGCAAATA	
11841	FORW	CATCACGTGACCTTCCAGTG	746
11841	REV	TCTCTCTCTCGCTCGGTTTC	

Supplementary Table 4. Primers used for ISH probe templates amplification. (*Continued*)

ID	Orient	Sequence 5'-->3'	Size
6217	FORW	TCAACAATGCCAATGATGCT	750
6217	REV	TCAGGCAGCATTTCAACAAG	
11284.1	FORW	CAGGCCAAGAAGAAGACACC	761
11284.1	REV	ACTGTTGATGCTGCAAGTCG	
7206	FORW	AAGGACGAAGACAGCGAGAA	746
7206	REV	AGCACGTAGACCTCCGGATA	
1323	FORW	GTGGTTAATCCGCATCTGCT	753
1323	REV	AACGCTGCCTGTTAGCTTGT	
4852.2	FORW	AGCTTCTGGACAGTGGAGGA	689
4852.2	REV	TCAGCCACGAACTCGATGTA	
2640	FORW	CGCAGTCAGCTTTTCACTTG	751
2640	REV	GAACTGGCGATTGTTGGACT	
6853	FORW	GGCGAGACTGCAGCAAAG	681
6853	REV	GGCAGGGGAGACTTAACACA	
2450.1	FORW	AATAGCAATGTCCGGCACTC	736
2450.1	REV	CGACATACCGTGGTGATTTG	
16357.1	FORW	ATCAGACGCAATATCGACCT	730
16357.1	REV	AAAAGGCGAATGACTGGATG	
2073	FORW	TCTGGTGCTCACCTGTCAAG	749
2073	REV	TGTCGTACTGCAAAGCGAAC	
3208	FORW	CATTCGTACGCTGTGCAACT	748
3208	REV	GAACACTTCGCCAAGGACTG	
13520.1	FORW	TTGCAAAAGGTGTTGCAGAA	746
13520.1	REV	ACGGCTGGCTTTTGATCTT	
1574.2	FORW	AGCGGAAGCTCGAACACAT	727
1574.2	REV	CGCAGCATTTGAACGAAATA	
9606	FORW	GAGCTGTGGTCATGGACCTT	713
9606	REV	ATCGCTCGTATGGTTAGACG	
18247.1	FORW	TCTATCGCGAGCTCAAAACTC	582
18247.1	REV	TCTATCCCGCTGTTTCTTCC	
27656.1	FORW	ATTCGGCCCAGTATCATCAA	626
27656.1	REV	TCTCTCGCCTGAGACAGATG	
10105	FORW	GCTGTCCAAGGTGCTTCTTC	750
10105	REV	ACCGTTCGATTCTCCACCTA	
13666.1	FORW	CTCGAGGGCCTGATCTTGTA	705
13666.1	REV	ACGTCGATGCCCGTATCC	
6652	FORW	CTTCAAGGTGGGCATCAACT	531
6652	REV	TTTAGCTGTGCGAAAAAGCA	
30114	FORW	AGCAGCAGCAGCAGTGG	644
30114	REV	TTCTCGTTGAAGGCAGCAG	
1534	FORW	ACTCCGGTGTCACCTTCAAC	749
1534	REV	GAGTCGATGCAGATGTCCAA	
5148	FORW	AGACTATCCGGCAGGCAATA	753
5148	REV	ATTCTCGCTGTCAACCGTCT	
2177	FORW	CAAGACGGAGATCTCGAAGC	754
2177	REV	AACAACTTGCGTCAGTGCAG	
2351	FORW	GTTTGGAACCCATCAGTGCT	754
2351	REV	ACAACCATAGCGGAAACCAA	
30797	FORW	CCCACACTTACAGAGTGCAAA	574
30797	REV	GCCAGCTGGAGATCCTTTAG	
5840	FORW	AAGTGCCGACTTCGTCTACC	746
5840	REV	CGGTTCAGTCGAAAGAGAGG	
108	FORW	AAAGTTCGTTGCACCAATCC	749
108	REV	TCCTCGCCCTGATAGACAGT	
5405	FORW	GAGGCCAAGAAGGTCAAGC	642
5405	REV	ATCCTGCGACTTGGTTTTGT	
30732	FORW	ACGGTGCCTCTGTGGAAATA	576
30732	REV	TCAAGCATCCAATTCAGCAG	
14439.3	FORW	CCTTCAAAGGGATGCTACGTT	414
14439.3	REV	GGAATAACCATTCCCCCAAT	

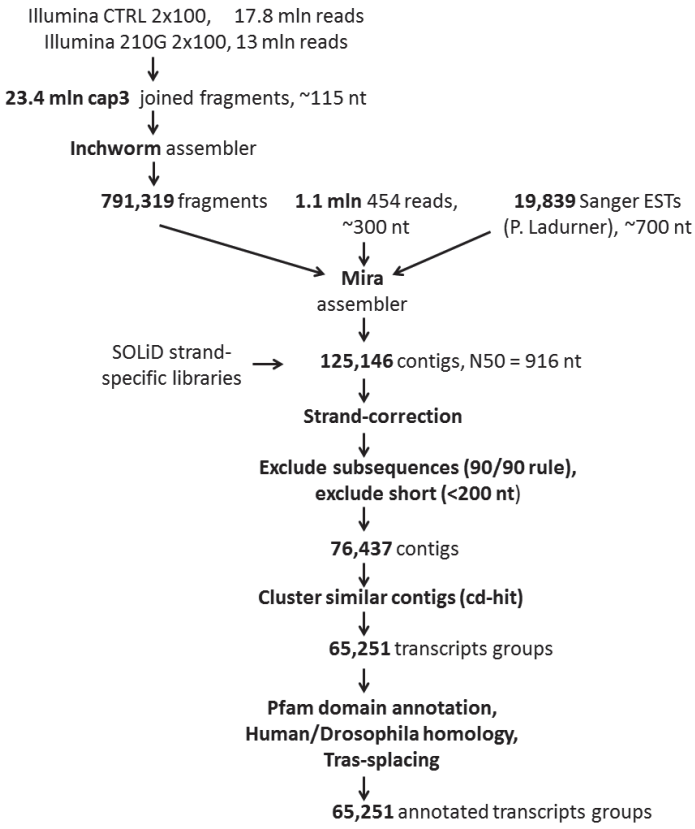
Supplementary Table 4. Primers used for ISH probe templates amplification. (*Continued*)

ID	Orient	Sequence 5'-->3'	Size
21543	FORW	GGTGAAGAGCTCGCAGATTT	674
21543	REV	CCGAGTGTGTGAAGCGTGAT	
30807	FORW	TTTTTGCGGAAAATTTGAAC	539
30807	REV	GTTGGCAAAGCCCTCAGTAG	
5707	FORW	GAGAGCACCCGGCTATACTG	748
5707	REV	TTCAGCTTCAAGGTGTGCAG	
10706.1	FORW	CTGGGAATGTCAATGCAGAA	742
10706.1	REV	AAAACAACGCACTGTGAAACC	
20011.2	FORW	AGAACGCTTGTTTCCATTGC	580
20011.2	REV	GAAGCCGAGTTCAACCTGTC	
7955.2	FORW	GGGATTCTACCAGCCTCAGC	712
7955.2	REV	GAGCAGAAGGCGGTCTTGTA	
2500.1	FORW	CGGGCATGGACTTCTACAGT	751
2500.1	REV	TTCTGCTGCACAGGTACTGG	
9892	FORW	GACTGCATCCGCTGTACAAA	742
9892	REV	CCGGTTCAATCGAACAAAGT	
12337.1	FORW	GGAATTCTGTCACCGACTGG	753
12337.1	REV	AATTTCTGCACCGTGCTTTT	
15645	FORW	ATGCTCAAGGACGCGACTAT	645
15645	REV	AGAGCCAGAAATGTCGGAAG	
36034	FORW	GCCAAGGCTGTGAAGTCTGT	538
36034	REV	ACAGTTCGCACCTGTGCTT	
25463	FORW	GGTCGGCATCATGAACAAG	641
25463	REV	AAAGCAGGAGTTCCAGAAACA	
10430	FORW	ATGGCTGTTAGCACCGCTAC	747
10430	REV	CCCAGTGCATCAAAAATGTG	
9315	FORW	TGCCAGATCCGGCTATTATC	735
9315	REV	ATTGGCGGCGATAGACATTA	
9951	FORW	AGCAACAGCAACAGCAACAG	747
9951	REV	CGCAGGTGTTCAAAGCTGTA	
18983	FORW	AAGATCCCAGATGCGACTTC	753
18983	REV	CTTTGTTCGCTTTGCTTTCC	
17297	FORW	GATCAAAACAACCGCACGTT	746
17297	REV	TCGGACGATAGGAAGGTGTC	
12957	FORW	GGGAGAATGCTACTCGGTGA	770
12957	REV	TCGCTGCAAACCAATAACTG	
13973	FORW	AAGCCCGACCTCTACAGCTT	720
13973	REV	GGTGAGGAACTCCTGCTGAG	
14005.1	FORW	CAATTGCTTGAAGCTTTCTGTG	670
14005.1	REV	CTATTGGTGGCCTGCTCATC	
6677	FORW	GACCTTTCTGCCGTGACCTA	751
6677	REV	AGCATGGCCAATGTCTAAGG	
11522.1	FORW	CTGGATTGACTCGGTCAGC	740
11522.1	REV	AATTCTACGAGCGCACCACT	
18995.1	FORW	TGAATCGAACCGATCAACAA	417
18995.1	REV	TAGATCTCGCTCGTGTCGAA	
12602	FORW	CACCAGCCGCTTATCAATG	708
12602	REV	ATTGCCGTTGAATGTGAGGT	
16202	FORW	CCCAAACTGTGGCATTGTG	659
16202	REV	GTACGTCGGACGGAACG	
9521.1	FORW	GTCTGGAGGAGGAGGAGGAC	692
9521.1	REV	TTGAAAGTAGAGCATCGCGTAA	
17643	FORW	AGTCAGCAGGGCTCAGACTC	754
17643	REV	ATCGTCACGTGAAGGGAAGT	
7259.2	FORW	ACCATAAAATGGCGTCAACC	713
7259.2	REV	GCTCTCCTTCGCAGACCTT	
17103.1	FORW	TGTCTGGCGACTACGAGGAT	659
17103.1	REV	TGAGAAGAGATGGAGGCAGAG	
59	FORW	CTGGCTGCGATCTGACAATA	753
59	REV	CTGCACCACTCTGATGCTGT	

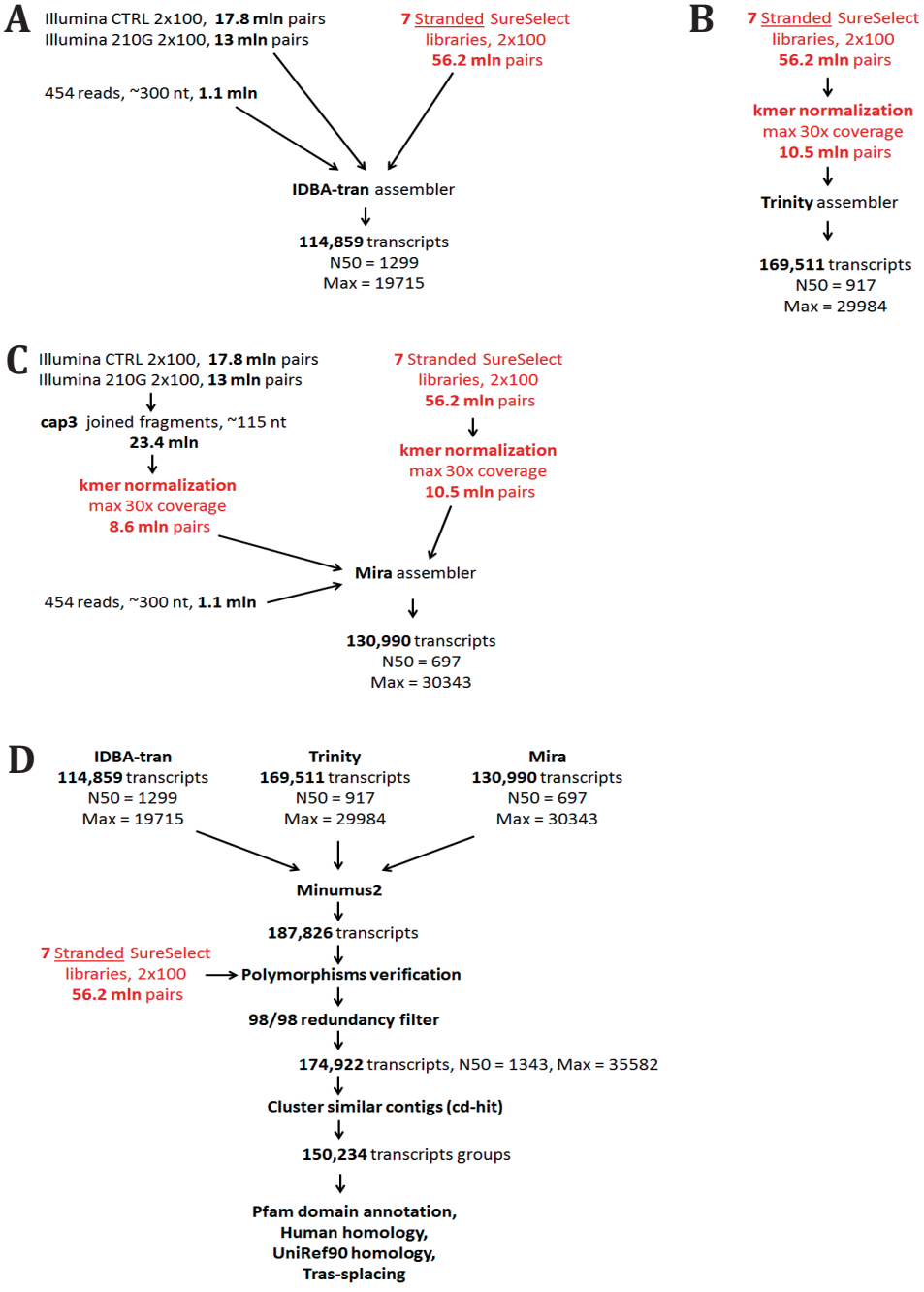
Supplementary Table 4. Primers used for ISH probe templates amplification. (*Continued*)

ID	Orient	Sequence 5'-->3'	Size
17489	FORW	TAGCATTCAAGCCAATGTGC	628
17489	REV	GCCGGTTACAAACAGACAGG	
18275	FORW	CGTCCGATATTGCTTGCTTT	751
18275	REV	GTTGAAAACGCCGAATTAGC	
3335	FORW	ACAATTGCCCGACTTCTTTG	753
3335	REV	CACACACTTGGTGGTGTGGT	
9816	FORW	GCCTTCATCAAGACCTGCAT	751
9816	REV	AGAGAGCTCAACTGGCGATT	
11498.1	FORW	ACTTGTCCAGAGCCGACCT	735
11498.1	REV	CGCTGTTGTTTCGGTGATTT	
3366.3	FORW	TTCCACACGCAAACAACTTC	726
3366.3	REV	TTTGCGTCTAGAGCTCAGGTT	
3838.3	FORW	TCATCAATGGACGCTTCATC	763
3838.3	REV	TTTACGGGACGGTAGTTCGT	
7117	FORW	GGTTGAGGTTGAAGCTACGC	743
7117	REV	GAAAAACCAGGGTTGGGACT	
31340	FORW	ATGACTATCATCCGCACGAC	615
31340	REV	AGTCAAGCAGACGCACATTG	
21261	FORW	CTCGGACTGCGGATACAAGT	686
21261	REV	AAGCTCACGAATTCTTTCTTGC	
20777.1	FORW	CGGGCTACTACTGCCTTCCT	400
20777.1	REV	AGTAGTTGCAGTGTGCATGG	
19389.1	FORW	CGGCGATAACTACGGCTCT	623
19389.1	REV	AGTGGTCCTGCGATTCACAC	
27244	FORW	ACTGTCAGCCTGGCCTGTT	560
27244	REV	AGTAAGACGGCCAGAACCAG	
20234	FORW	GGGACATCGCGGTAGAGAAG	474
20234	REV	CGAGCATGCAAATAGCAAGA	
17987	FORW	CGGAATCGGAGACAGGTTTA	703
17987	REV	TCTAGCCCTTTTTCCTCGAA	
3837	FORW	TGAGATCAAGTCTAGCCCCAAT	492
3837	REV	TCAAAACGAAATTACGACATGG	
369.3	FORW	ATTCGCTGACCAGCAAGAGT	510
369.3	REV	GGGGAAAGAGTTTCGGTACA	
3584.1	FORW	TCGACTCTTAGGACGCACCT	755
3584.1	REV	TGCGGGTACCACTAGACACA	
12596	FORW	GGAGCTCTACAGCGACTGGT	755
12596	REV	CGAGTTGGGTTCGCAATAAT	
27827	FORW	ACCACTACTGCCGCTACCAG	563
27827	REV	ATAAACTCGTCTGCGGGTTG	

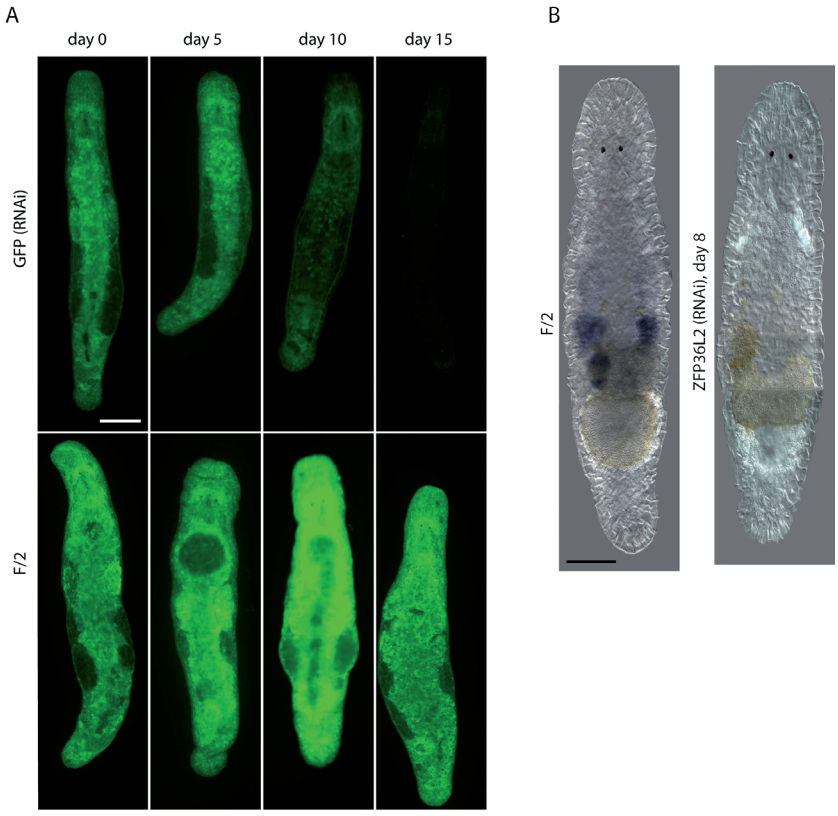
ID – number of the transcriptome in MLRNA110815 assembly; Orient – primer orientation (FORW – forward; REV – reverse); Size – size of the PCR product, bp.



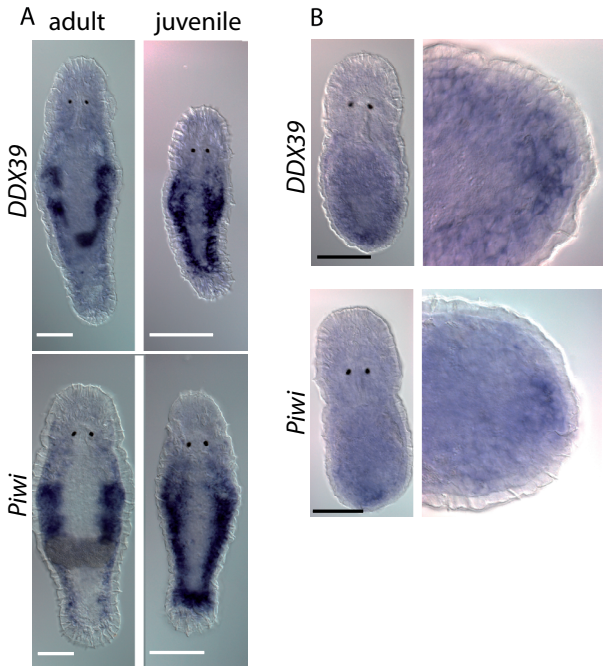
Supplementary Figure 1. Steps for generating MLRNA110815 *de novo* transcriptome assembly. See text for details.



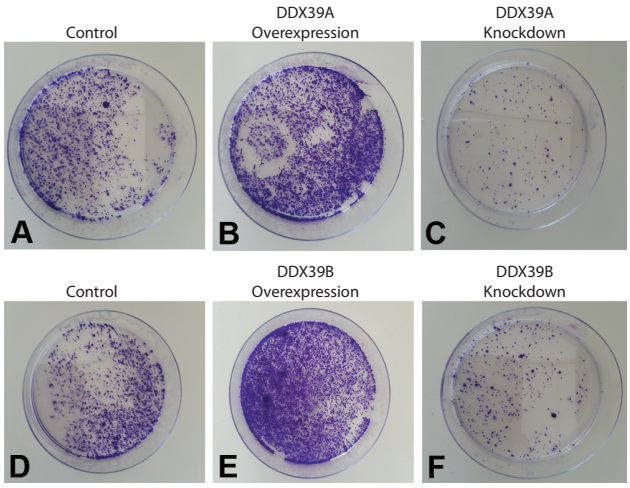
Supplementary figure 2. Steps for generating MLRNA131024 *de novo* transcriptome assembly. (A) IDBA-tran assembly using 454 data and both stranded and non-stranded Illumina data. (B) Trinity assembly using stranded Illumina data normalized to 30x kmer coverage. (C) Mira assembly using Illumina kmer coverage normalized data and 454 data. (D) Merging of the three assemblies with Minumus2 and downstream processing of transcripts. See text for details.



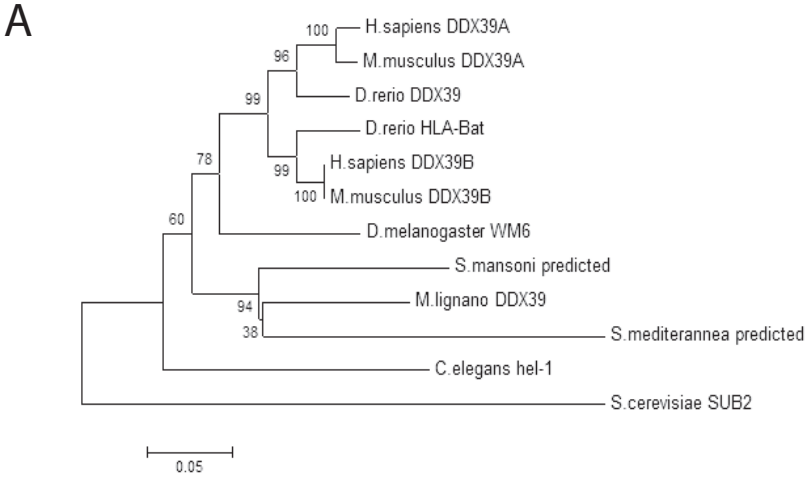
Supplementary Figure 3. RNAi-mediated knockdown efficiency. (A) Transgenic Hub1 worms (expressing GFP) incubated in f/2 and GFP dsRNA solution (30ng/µl) at different time points after start of the treatment. After 2 weeks of dsRNA treatment no GFP expression can be detected. Pictures were taken with a Leica DM 6000 microscope with same exposure time and intensity settings within every time point. (B) Expression of *ZFP36L2* gene in the control worm and after 9 days of dsRNA treatment checked by ISH. No expression can be detected after RNAi treatment. Scalebar 100 µm (all worms).

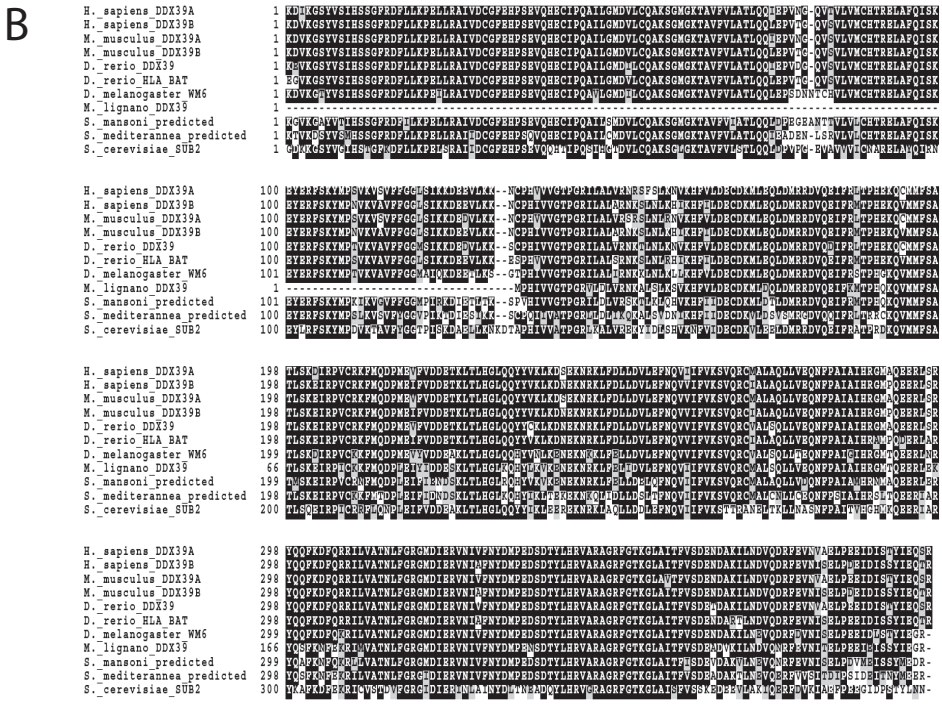


Supplementary Figure 4. *DDX39* expression in adult and juvenile worms in comparison with the expression of neoblast-marker *Piwi*. (A) Both genes are expressed in gonads and somatic stem cells. Expression in the neoblasts is obvious in juvenile worms, that do not have gonads yet. (B) Expression of both genes is upregulated in blastema at the wound site 12 hours after cutting. Scalebar 100 μm (all worms).



Supplementary Figure 5. Colony formation rate in cells depends on DDX39 expression. (A,D) Control plates after 10 days of seeding the cells. (B) Overexpression of DDX39A accelerates formation of colonies, whereas knock-down of DDX39A (C) block proliferation. Similar effects are observed for DDX39B overexpression (E) and knockdown (F).





Supplementary Figure 6. DDX39 gene is conserved in evolution. (A) Neighbor-Joining tree of *DDX39*. (B) *DDX39* protein sequence alignment in vertebrate and invertebrate model species. Species abbreviations: *S. cerevisiae, Saccharomyces cerevisiae; C. elegans*, *Caenorhabditis elegans; S. mediterranea, Schmidtea mediterrane; M. lignano*, *Macrostomum lignano*; *S. mansoni, Schistosoma mansoni; D. melanogaster, Drosophila melanogaster*; *M. musculus*, *Mus musculus*; *H. sapiens*, *Homo sapiens*; *D. rerio, Danio rerio.*

CHAPTER 4

THE FLATWORM *MACROSTOMUM LIGNANO* IS A POWERFUL MODEL ORGANISM FOR ION CHANNEL AND STEM CELL RESEARCH

Daniil Simanov[1], Imre Mellaart-Straver[1], Irina Sormacheva[2] and Eugene Berezikov[1,3]

[1]Hubrecht Institute, Hubrecht Institute, KNAW and University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands
[2]Institute of Cytology and Genetics SB RAS, 630090 Novosibirsk, Russia
[3]European Research Institute for the Biology of Ageing, University of Groningen, University Medical Centre Groningen, 9713AV Groningen, The Netherlands

Stem Cells International (2012): 167265.

ABSTRACT

Bioelectrical signals generated by ion channels play crucial roles in many cellular processes in both excitable and non-excitable cells. Some ion channels are directly implemented in chemical signaling pathways, the others are involved in regulation of cytoplasmic or vesicular ion concentrations, pH, cell volume and membrane potentials. Together with ion transporters and gap junction complexes ion channels form steady-state voltage gradients across the cell membranes in non-excitable cells. These membrane potentials are involved in regulation of such processes as migration guidance, cell proliferation and body axis patterning during development and regeneration. While the importance of membrane potential in stem cell maintenance, proliferation and differentiation is evident, the mechanisms of this bioelectric control of stem cell activity are still not well understood, and the role of specific ion channels in these processes remains unclear. Here we introduce the flatworm *Macrostomum lignano* as a versatile model organism for addressing these topics. We discuss biological and experimental properties of *M. lignano*, provide an overview of the recently developed experimental tools for this animal model, and demonstrate how manipulation of membrane potential influences regeneration.

1. INTRODUCTION

Ion channels represent a diverse family of pore-forming proteins. They are crucial for establishing voltage gradients across plasma membranes by allowing the flow of inorganic ions (such as Na^+, K^+, Ca^{2+} or Cl^-) down their electrochemical gradients. Ionic flux through the channels provides the foundation for membrane excitability, which is essential for the proper functioning of neurons, cardiac and muscle cells (Hille, 2001). At the same time, ion channels serve many functions apart from electrical signal transduction. For example, Ca^{2+} is an important messenger, and changes in its intracellular concentrations influence numerous cellular processes in virtually all types of non-excitable cells (Fewtrell, 1993; Thomas et al., 1996; Schuster et al., 2002), including stem cells (Tonelli et al., 2012; Apáti et al., 2012; Paredes-Gamero et al., 2012). Besides, a number of ion channels are known to be directly involved in chemical signaling pathways in different cell types (Sheng and Pak, 2000; Arcangeli and Becchetti, 2006). As a result, mutations in genes encoding ion channel proteins have been associated with many disorders (so-called "channelopathies"), caused by dysfunction of both excitable (epilepsy, hypertension, cardiac arrhythmia) and non-excitable (diabetes, osteopetrosis, cystic fibrosis) cells (Hübner and Jentsch, 2002). Here we briefly describe the crucial role ion channels play in maintenance, proliferation and differentiation of stem cells on the level of single cell and the whole organism. We discuss the importance of animal model systems, such as flatworms, for studying bioelectric signaling in complex morphogenesis during development and regeneration. Finally, we introduce the new flatworm model, *Macrostomum lignano*, and discuss its potential to extend our knowledge on the subject.

2. ION CHANNELS AND MEMBRANE POTENTIAL IN STEM CELLS

Numerous ion channels and pumps together with gap junction complexes form transmembrane voltage gradients. While quick changes of these membrane potentials (V_{mem}) are best described in neurons, muscle and cardiac cells, long-term steady-state V_{mem} levels are present in all other cells (Pandiella et al., 1989; Lang et al., 1991). Membrane potentials strongly correlate with the mitotic ability of different cell types, with the high resting potential associated with differentiated non-dividing cells (Binggeli and Weinstein, 1986). V_{mem} fluctuations during progression through the cell cycle have been reported in a number of cell types, and changes of membrane potential appears to be required for both G1/S and G2/S phase transitions (Cone, 1969; Freedman et al., 1992; Blackiston et al., 2009). Modulation of V_{mem} through applied electric fields or by inhibition of ion channels leads to cell cycle arrest in dividing cells (DeCoursey et al., 1984; Chiu and Wilson, 1989; Amigorena et al., 1990; Wang et al., 2003), and artificial membrane hyperpolarization induces differentiation of mesenchymal stem cells (Sundelacruz et al., 2008). On the other hand, electroporation (supposedly followed by membrane depolarization) activates cell hyperproliferation and de-differentiation (Atkinson et al., 2006).

On the level of multicellular organism, progression through the cell cycle should be strictly regulated and synchronized during such processes as development and regeneration in order to achieve a proper body patterning. Accordingly, stable and reproducible membrane

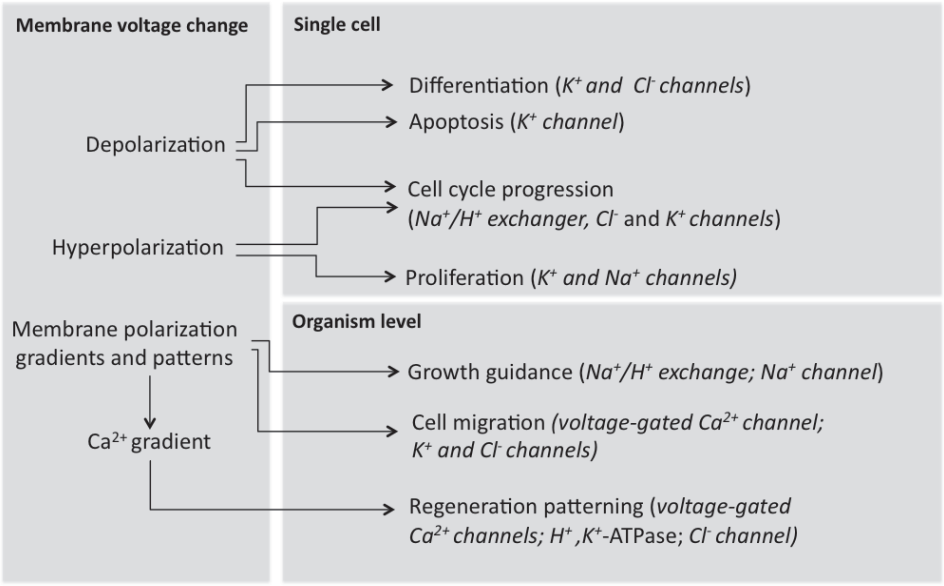


Figure 1. Ion channels and membrane voltage during regeneration. Changes of membrane potentials can directly affect different aspects of cell behavior and large-scale morphogenetic processes during regeneration. Ion channels and transporters implicated in these processes are mentioned in brackets.

polarization patterns have been recently described in various model organisms. Artificial modulation of these patterns during development or regeneration has a large impact on left-right asymmetry and anterior-posterior identity (Adams et al., 2006; Adams et al., 2007; Nogi et al., 2009; Beane et al., 2011; Zhang et al., 2011). The role of bioelectric signaling in regeneration is comprehensively reviewed in (Levin, 2009) and schematically shown in Figure 1. Finally, modulations of membrane voltage have been observed in a large number of oncologic disorders, and ion channels were proposed as cancer treatment targets (Kunzelmann, 2005; Arcangeli et al., 2009).

Thus, bioelectric signaling is an important mechanism of cell regulation, including stem cell maintenance, proliferation and differentiation. Recent findings suggest this control system to be well-conserved in a wide range of animal phyla. However, the mechanisms linking membrane potential to the cell cycle, proliferation and differentiation, and the role of specific ion channels in this process remain largely unclear. The picture becomes even more complicated on the level of multicellular organism. Our understanding of the ways cells produce and receive bioelectric signals and translate them into positional information during development and regeneration is still fairly poor. While considerable knowledge about the role of membrane potential in stem cells was gathered recently from different species, the number of models used in this field is still limited. Expanding the range of model organisms used for functional studies of bioelectric signaling is crucial for better understanding of this control system and its role in complex morphogenesis.

3. PLANARIAN MODELS IN ION CHANNEL RESEARCH

Planarian flatworms are long-established models for stem cell and regeneration research. The adult stem cell system and regeneration capacity of the species *Planaria maculata* and *Planaria lugubris* were described by Thomas Hunt Morgan as early as in the end of 19th century (Morgan, 1898; Morgan, 1902). In our days the favorite planarian species in the regeneration field are *Schmidtea mediterranea* and *Dugesia japonica* (Sánchez Alvarado, 2004; Agata et al., 2006).

Planaria were also one of the first species in which stable membrane potential patterns were described, and their role in regeneration postulated. In 1940s and 1950s Marsh and Beams were able to specifically control establishing of anterior-posterior axis by providing bioelectrical signals to regenerating planaria fragments (Marsh and Beams, 1947; Marsh and Beams, 1952; Dimmitt and Marsh, 1952).

In the last 5 years considerable work was done in planaria on understanding the molecular and genetic mechanisms that allow cells to establish and maintain long-term membrane potential patterns and transduce bioelectric signals into proliferation and differentiation decisions. The importance of gap junction signaling in establishing anterior-posterior polarity during regeneration was shown (Nogi and Levin, 2005), and the specific innexin gene, *Smedinx-11*, responsible for blastema (regenerating tissue) formation and stem cell maintenance identified (Oviedo and Levin, 2007).

The role of ion channels and pumps in establishing of anterior-posterior axis during regeneration of planaria *D. japonica* was recently highlighted by groups of Michael Levin and Jonathan Marchant. *D. japonica*, which can regenerate an entire animal from a small part of a cut worm, has highly depolarized cell membranes in the head region, and highly polarized – in the posterior part. In the cut worm this pattern is re-established rapidly, regardless of the cutting plane (Beane et al., 2011). After the wound is closed, blastema at all anterior-facing wounds gives origin to heads, while tails are regenerated from the posterior-facing wounds. The polarization pattern was altered by highly specific drugs against different ion channels and transporters, such as SCH-28080 (inhibitor of H^+,K^+-ATPase), ivermectin (IVM, activator of the invertebrate GluCl channels) or praziquantel (PZQ, activator of voltage-operated Ca^{2+}-channels). Remarkably, induced depolarization itself is sufficient to drive ectopic anterior (head) regeneration even in posterior-facing blastemas, whereas membrane polarization of anterior-facing wounds blocks the head regeneration (Nogi et al., 2009; Beane et al., 2011). The role of specific voltage-operated Ca^{2+}-channels in regenerative patterning was addressed in the follow-up experiments (Zhang et al., 2011).

Thus, planarian flatworms can be successfully used for ion channel and stem cell studies. Fascinating regeneration capacity of these animals, together with a wide range of research techniques established and optimized over the last 100 years make planaria a very attractive model for studying bioelectric signaling during regenerative morphogenesis. However, due to inefficient sexual reproduction under laboratory conditions, classical genetic methods are not available in planarian flatworms, and reverse genetics methods are limited to RNA interference. Since genetic manipulation of these animals is difficult, no reproducible transgenesis methods are available for planaria (Poss, 2010).

4. EXPERIMENTAL PROPERTIES OF THE FLATWORM *MACROSTOMUM LIGNANO*

During the last decade another flatworm, *Macrostomum lignano*, has emerged as a complementary model organism for regeneration research (Ladurner et al., 2005; Egger et al., 2006; Pfister et al., 2007; De Mulder et al., 2009). This marine free-living flatworm is about 1,5 mm long and consists of roughly 25000 cells. *M. lignano* is easy to culture in laboratory conditions, and populations of this animal are continuously maintained in the number of laboratories for over a decade. The generation time of the flatworm is short, with about two weeks of postembryonic development to sexually mature adult. Both juvenile and adult worms have clear morphology and are highly transparent (Fig. 2A), greatly facilitating phenotyping and both fluorescent and non-fluorescent staining. The regeneration capacity of *M. lignano* is provided by roughly 1600 neoblasts (adult stem cells) located mesodermally. Proliferation activity of these cells can be easily studied using BrdU labeling, performed by simple soaking (Pfister et al., 2007; Pfister et al., 2008) (Fig. 2B). Importantly *M. lignano* is non-self fertilizing hermaphrodite and has exclusively sexual reproduction. Well-fed adult animal generate a lot of embryos all year through (one animal lays one egg a day on average), making it accessible for genetic manipulation. *In situ* hybridization (Pfister et al., 2007) and RNA interference (by soaking) (Pfister et al., 2008) protocols are established and optimized for *M. lignano*, and a number of tissue-specific monoclonal antibodies are available (Ladurner et al., 2005). Basic culturing and experimental properties of *M. lignano* are summarized in Table 1.

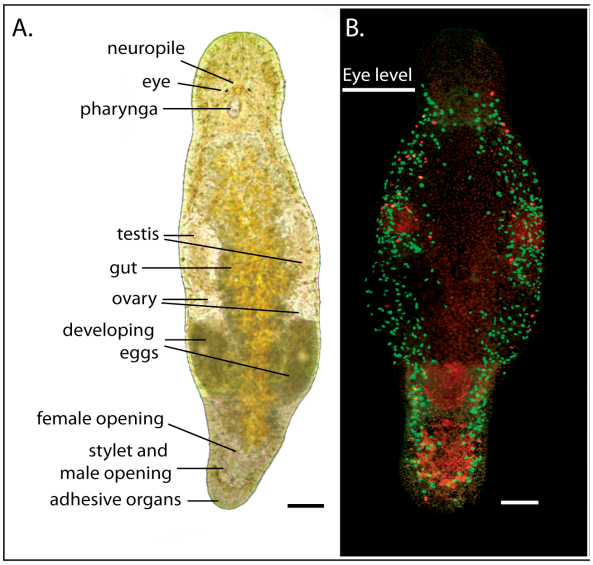


Figure 2. *Macrostomum lignano* as a model organism. **(A)** Bright field image of a living specimen. **(B)** Confocal projection of BrdU and phospho- histone H3 immunostaining after 30 minutes BrdU pulse in an adult worm (green: S-phase cells, red: mitotic cells). Scalebar 100 mm.

Table 1. Culturing and experimental properties of *M. lignano*.

Size	1 mm
Total cell number	+/- 25.000
Neoblasts	+/- 1600
Transparency	Highly transparent
Culturing media	f/2 (sea water based)
Feeding	Diatom algae (*Nitzschia curvilineata*)
Embryogenesis	5 days
Generation Time	18 days
Nervous-, muscle system and gonads	Simple
Stem Cell System	Pluripotent
BrdU/H3 Staining	Yes (easy by soaking)
RNA interference	Yes (easy by soaking)
Accessibility to eggs	Single eggs (one egg/day)
Transgenics	Possible, by injection into eggs

During the past three years considerable progress has been made towards establishing *M. lignano* as versatile stem cell research model for the genomics era. The work on *M. lignano* genome assembly and annotation is in progress (Simanov et al., in progress), and draft genome and transcriptome assemblies are publicly available at http://www.macgenome.org. Comparing transcriptome data obtained from irradiated (neoblast-depleted) and control worms provided the insight into the role of a number of genes in regeneration, while stage-specific transcriptome data showed the temporal expression of *Macrostomum* genes through development (Simanov et al., in preparation). Most importantly, proof-of-principle for transgenesis in *M. lignano* has been demonstrated and first stable transgenic GFP-expressing lines of *M. lignano* are established (Demircan, De Mulder, Berezikov et al, in preparation). Thus, biological and experimental properties of *M. lignano*, combined with its rapidly expanding experimental toolbox, make this animal an attractive and powerful model organism for stem cell and regeneration research. Its astonishing ability to resist γ-irradiation and recover after being exposed to it makes the neoblast system of this animal exceptional even for flatworms (De Mulder et al., 2010). Moreover, fascinating but yet poorly understood link between regeneration and rejuvenation provides exciting opportunity of using *M. lignano* as a model for ageing research (Mouton et al., 2009).

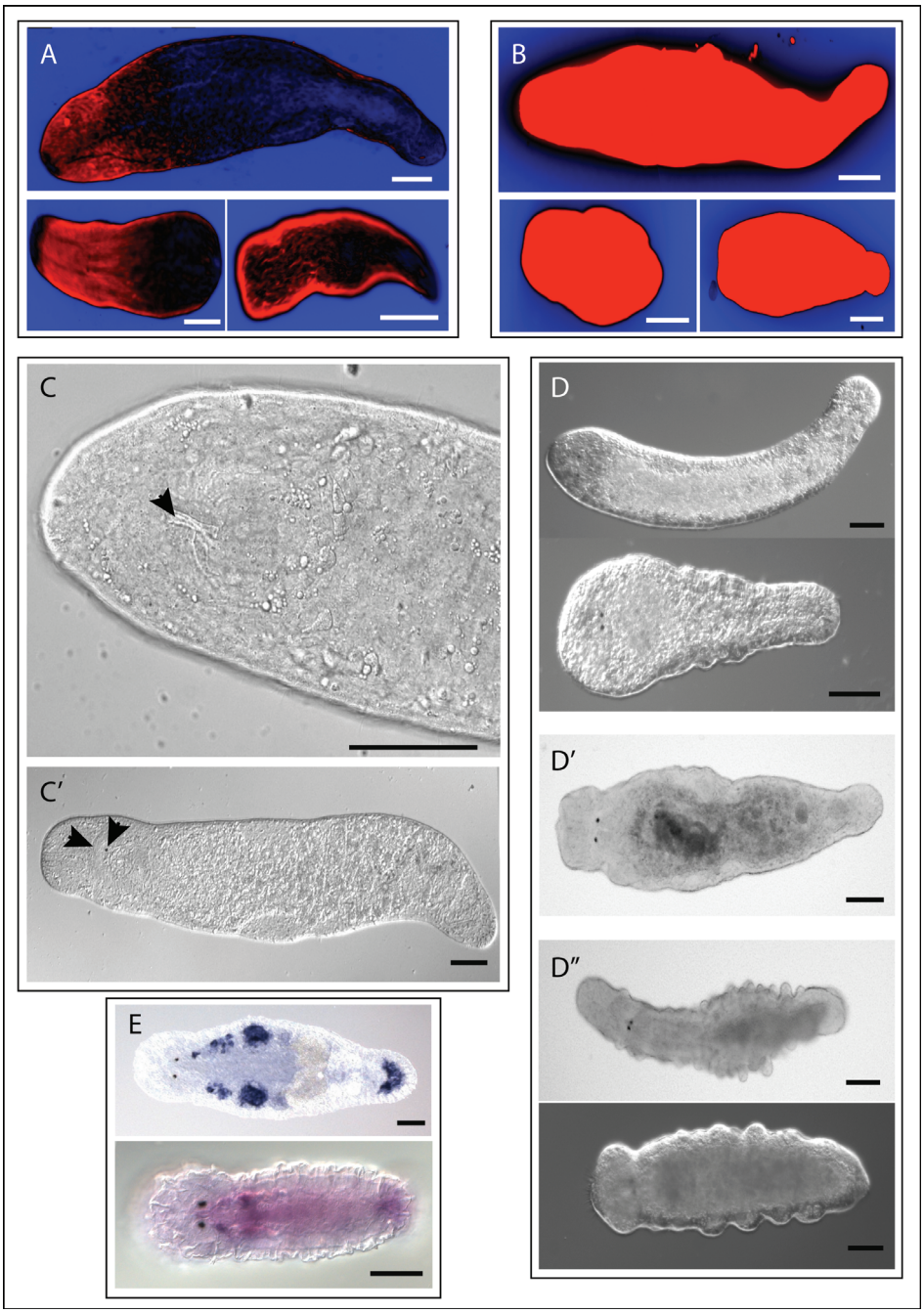
5. ION CHANNELS AND REGENERATION IN *M. LIGNANO*

Unlike planarian flatworm species, *M. lignano* is unable to regenerate the head under normal circumstances. Posterior-facing blastemas give origin to fully functioning tails with all its organs and structures, whereas anterior-facing wounds develop blastema layer but the actual

regeneration can only happen if the worm was amputated in front of the brain (at the very tip of the head). Thus, anterior fragments of the worm, having a functional head, can regenerate the whole body in 2-3 weeks, while posterior fragments normally die 5-10 days after losing the head (Egger et al., 2006). These differences in the head regeneration capacity between *M. lignano* and planarians, and the ability to induce ectopic head regeneration in *D. japonica* by the manipulation of membrane voltage gradients, prompted us to investigate how these findings in planarians translate into *M. lignano*. $DiBAC_4(3)$ voltage-reporting dye stainings (as described in (Oviedo et al., 2008)) showed that membrane voltage pattern in *M. lignano* is similar to the one observed in *D. japonica* (Beane et al., 2011) – the anterior part is highly depolarized, while the tail is relatively polarized. In the cut worms this pattern is quickly re-established in the anterior head-containing fragments, while the posterior headless fragments do not show any clear anterior-posterior polarization gradient and do not regenerate (Fig. 3A). Just like in planarian flatworms, membrane polarization patterns in *M.* lignano can be altered using drugs against ion channels. IVM induces depolarization of the membranes of intact and cut worms, both in anterior and posterior regions (Fig. 3B). Posterior-facing blastemas still regenerate the tails after treatment, though the full regeneration takes longer than normally. Anterior-facing wounds treated with IVM develop blastema, and some tissue growth is often observed within a week after wound closure. IVM-treated headless fragments always move more actively and survive longer comparing to control fragments. Strikingly, 1,5% of posterior fragments after IVM treatment are able to regenerate head-specific structures and, in a few cases, a fully functional head (Fig. 3C, C'). PZQ causes the same depolarization effect but does not have any effect on regeneration patterning at tested concentrations (data not shown). Intact animals exposed to high doses of IVM or PZQ display phenotypes that in planarian flatworms are stereotypically associated with stem cell loss or disorder (Reddien et al., 2005; Guo et al., 2006; Pearson and Sánchez Alvarado, 2010; Cowles et al., 2012). *M. lignano* animals treated with 2μM IVM gradually lose anterior identity, with no head-specific structures left 7-9 days after treatment (Fig. 3D). After exposure to higher doses of IVM (3-4 μM) worms develop characteristic square head due to partial tissue loss in the most anterior part of the body, get paralyzed, and die 3-4 days after treatment (Fig. 3D'). High concentration of PZQ in culturing media causes formation of bulges, mainly in the posterior part of the body (Fig. 3D''). This phenotype is completely different from the one observed after IVM treatment, suggesting specific action of the drugs.

These pilot experimental results show that *M. lignano* can be successfully used as a model for ion channel and stem cell studies. The complete transcriptome and established *in situ* hybridization and RNA interference methods, in combination with chemical treatment make

Figure 3 (next page). Bioelectric signaling and stem cells in *M. lignano*. (A-B) $DiBAC_4(3)$ staining of intact worm (top), anterior (left bottom) and posterior (right bottom) fragments. **(A)** – control worm, **(B)** – worm treated with 1μM IVM. Blue is more polarized than black, black is more polarized than red. (C-C') Regeneration of head-specific structures after 1 μM IVM treatment. Arrowheads in **(C)** indicate regenerated pharynx, in (C') – regenerated eye and half of the brain. (D-D'') intact worms exposed to high doses of IVM (2 μM in D and 4μM in D') and PZQ (150 μM in D''). **(D)** – head regression; (D') – square head; (D'') – bulges and outgrowth. **(E)** *In situ* hybridization results in adult (top) and juvenile (bottom) animals with the probe against RNA815_5834 transcript from ML110815 transcriptome assembly (voltage-gated sodium channel). In juvenile worm this gene ▶



► is expressed almost ubiquitously, and in adults expression is only detected in gonads and (likely) in somatic stem cells. Strong signal in the adhesive glands in the tail is likely a common artifact.

it possible to address the function of specific ion channels in development, tissue turnover and regeneration. For example, comparison of transcriptome data from irradiated (stem cell-deficient) and non-irradiated animals highlights a number of ion channel genes expressed specifically in dividing cells (Fig. 3E), and future elaborated studies of such genes may provide novel insight into the role of bioelectric signaling in stem cell maintenance and differentiation. Importantly, a significant number of ion channels are well-conserved between *M. lignano* and human (Table 2), increasing the relevance of findings in flatworms to understanding ion channels and stem cells in human situation.

Table 2. Major categories of ion channel genes conserved between *H. sapiens* and *M. lignano*.

GO term	Description	H	M	Human genes
GO:0004889	acetylcholine-activated cation-selective channel activity	13	132	CHRNA4, CHRNE, CHRNA10, CHRNB1, CHRNB3, CHRNA6, CHRNA3, CHRND, CHRNB2, CHRNB4, CHRNA9, CHRNA2, CHRNA7
GO:0004931	extracellular ATP-gated cation channel activity	5	15	P2RX6, P2RX7, P2RX5, P2RX4, P2RX2
GO:0004970	ionotropic glutamate receptor activity	12	67	GRIN1, GRIA4, GRIN2A, GRIK2, GRIK1, GRIA1, GRIK4, GRIA2, GRIK3, GRID1, GRIN3A, GRIK5
GO:0005216	ion channel activity	6	23	PKD1L2, MCOLN3, MCOLN2, PKD2L2, PKD2L1, PKDREJ
GO:0005221	intracellular cyclic nucleotide activated cation channel activity	2	5	KCNA10, CNGA3
GO:0005222	intracellular cAMP activated cation channel activity	1	2	HCN4
GO:0005223	intracellular cGMP activated cation channel activity	1	1	CNGB3
GO:0005229	intracellular calcium activated chloride channel activity	2	3	ANO1, ANO2
GO:0005232	serotonin-activated cation-selective channel activity	2	3	HTR3B, HTR3A
GO:0005237	inhibitory extracellular ligand-gated ion channel activity	2	3	GABRA6, GABRB2
GO:0005242	inward rectifier potassium channel activity	6	26	KCNH6, KCNJ12, KCNK6, KCNJ8, KCNQ5, KCNH7
GO:0005245	voltage-gated calcium channel activity	8	27	CACNA1C, CATSPER1, CACNG7, CACNG5, CACNB1, CACNA1B, CACNB2, CACNA1E
GO:0005247	voltage-gated chloride channel activity	6	14	CLCN7, CLCN4, CLIC1, CLIC4, CLIC6, CLCN3
GO:0005248	voltage-gated sodium channel activity	8	19	SCN3A, SCN2A, SCN4A, PKD2, SCN8A, SCN5A, SCN9A, SCN11A
GO:0005249	voltage-gated potassium channel activity	23	75	KCTD12, KCTD21, KCNH3, KCTD10, KCTD3, KCTD6, KCNAB3, KCTD2, KCTD15, KCTD7, KCNH4, KCNB1, KCTD9, KCNH8, KCNC3, KCNC2, KCTD16, KCND1, KCNC1, KCNV2, KCNH5, KCTD1, KCTD20

Table 2. Major categories of ion channel genes conserved between *H. sapiens* and *M. lignano*. (*Continued*)

GO term	Description	H	M	Human genes
GO:0005250	A-type (transient outward) potassium channel activity	3	11	KCNIP2, KCND3, KCND2
GO:0005251	delayed rectifier potassium channel activity	8	26	KCNA3, KCNB2, KCNH2, KCNA1, KCNA5, KCNQ1, KCNA2, KCNH1
GO:0005254	chloride channel activity	17	55	CLCA1, ANO3, GABRB3, GABRA2, GABRB1, ANO7, ANO9, ANO4, GABRG2, CLCA4, CLCC1, ANO6, GABRQ, GABRG1, ANO10, GABRA4, GABRG3
GO:0005261	cation channel activity	7	33	TRPM3, TRPV4, TRPM6, TRPC7, TMEM38A, TRPV1, HCN2
GO:0005262	calcium channel activity	7	51	TRPM1, TRPM7, TRPM8, TRPV5, TRPM5, TRPM4, TRPV6
GO:0005267	potassium channel activity	13	27	KCNC4, KCNK16, KCNK10, KCNG1, KCNK2, KCNK5, KCNK3, KCNK12, KCNQ4, KCNK17, KCNIP1, KCNIP4, KCNK9
GO:0005272	sodium channel activity	4	40	HCN1, NALCN, ACCN4, TRPM2
GO:0008308	voltage-gated anion channel activity	2	3	VDAC1, VDAC2
GO:0008331	high voltage-gated calcium channel activity	7	38	CACNA1A, CACNA2D4, CACNA1D, CACNA2D1, CACNA1S, CACNA2D3, CACNA2D2
GO:0008332	low voltage-gated calcium channel activity	3	18	CACNA1H, CACNA1I, CACNA1G
GO:0015269	calcium-activated potassium channel activity	9	49	KCNMA1, KCNN1, KCNT2, KCNN2, KCNT1, KCNU1, KCNMB2, KCNK18, KCNN3
GO:0015276	ligand-gated ion channel activity	2	4	CLCA2, CNGB1
GO:0015279	store-operated calcium channel activity	5	40	TRPC4, TRPC6, ORAI1, TRPA1, TRPC3
GO:0015280	ligand-gated sodium channel activity	8	97	SCNN1B, SCNN1G, ACCN1, ACCN3, SCNN1A, ACCN5, ACCN2, SCNN1D
GO:0022824	transmitter-gated ion channel activity	4	34	GLRA2, GLRA4, GLRA1, GLRA3
GO:0030171	voltage-gated proton channel activity	1	3	HVCN1
GO:0072345	NAADP-sensitive calcium-release channel activity	2	3	TPCN1, TPCN2
	Total	199	947	
	Total number of genes in these GO categories	390		

H – number of different ion channel genes in human with homologs in *M. lignano*. M – number of transcripts in *M. lignano de novo* transcriptome assembly ML110815 with homology to ion channel genes in human. Note that alternatively spliced transcripts are counted separately in the *M. lignano* transcriptome assembly, hence the total reported number of transcripts is higher than the number of corresponding human genes. For this classification, genes were assigned to the least frequent available GO term within pre-defined list of ion channel-related GO terms (molecular function domain).

6. FUTURE DIRECTIONS

Macrostomum lignano has a great potential as a model for ion channel and stem cell research. The genetic toolbox available for this organism is already useful enough to address a wide range of scientific problems, and more methods and approaches will be optimized and used in this flatworm in the near future. *M. lignano* is a small animal and it is cultured in water, which makes it easy to apply different chemicals to the worms. Another major advantage of the animal is its high transparency. Phenotypic changes, fluorescent signals or certain transgene expression can be observed in any part of the body, as well is on the whole organism scale. For example, various fluorescent reporter dyes can be just added to culturing media in order to enable real-time *in vivo* monitoring of membrane potentials, pH and ion flows (Wolff et al., 2003). Short generation time and efficient reproduction of *M. lignano* make logistics of large-scale experiments, such as drug screens, feasible in this animal.

As a model, *M. lignano* offers an exciting opportunity to bridge the gap between bioelectric signaling and genetic pathways involved in stem cell functions. The expression pattern and

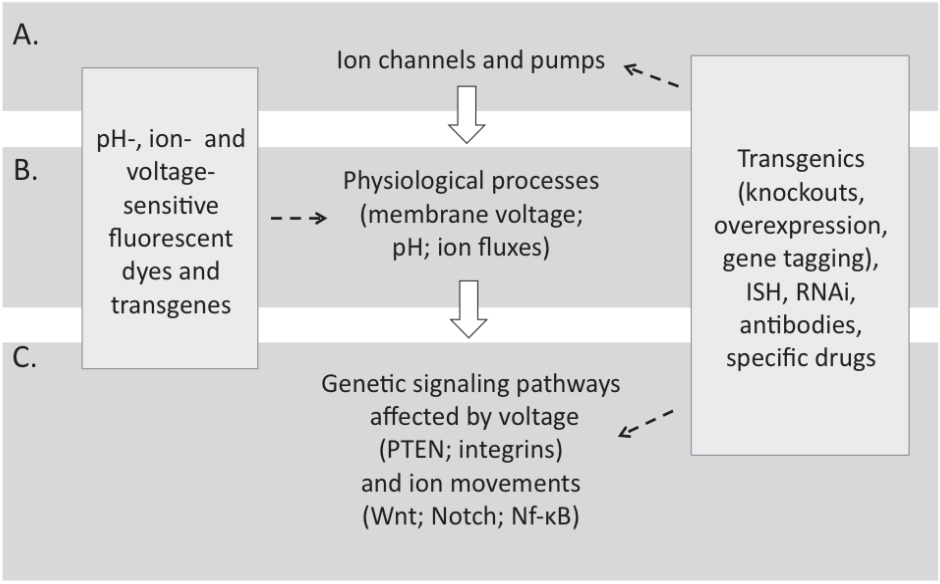


Figure 4. Studying regulation of stem cells in *M. lignano* on both genetic and epigenetic level. **(A)** Expression, localization and function of ion channels and pumps that give rise to bioelectric signals can be addressed in *M. lignano* by established methods such as RNAi or ISH in combination with specific drugs, antibodies and transgenics. **(B)** Changes in ion flows, pH and membrane voltage caused by these channels and pumps can be detected with sensitive fluorescent dyes or followed *in vivo* in mutants expressing pH-, or ion-sensitive forms of fluorescent proteins. **(C)** These processes affect known (and possibly unknown) genetic signaling pathways via different mechanisms including changes of Ca^{2+} concentrations, voltage-sensing domains of proteins and voltage-gated transport of signaling molecules. These pathways and functional link between genetic and epigenetic mechanisms of stem cell function regulation can be studied in transgenic mutant lines with the help of RNAi and ISH techniques.

function of any gene can be determined by *in situ* hybridization and RNAi protocols, but it is transgenics that can bring such studies to the whole new level. Transgenic reporter lines expressing pH-sensitive or Ca^{2+}-sensitive fluorescent proteins (Mahon, 2011; Zhao et al., 2011) would make a perfect tool to visualize bioelectric phenotypes during drug- or RNAi-screens. Overexpression of ion channels or even certain subunits would help to better understand their functions and interactions.

Targeted genome editing by Zinc Finger Nucleases have not been tested yet in this animal but should be also feasible and potentially can be used to generate ion channel knockout and knock-in lines (Liu et al., 2010; Wang et al., 2012). The same method allows fluorescent tagging of genes of interest and analysis of their expression, localization and functions at the endogenous level (Doyon et al., 2011). Sexual reproduction and lack of self-fertilization makes possible crossing different lines of *M. lignano* and hence to use the power of classical genetics approaches in this animal. Taken all together, we are convinced that *M. lignano* is poised to become a productive model to study relations between ion channels and stem cell regulation (Fig. 4).

ACKNOWLEDGEMENTS

We thank S. Mouton for providing the image of the BrdU-labeled *M. lignano* animal.

REFERENCES

Adams, D.S., Masi, A., and Levin, M. (2007). H+ pump-dependent changes in membrane voltage are an early mechanism necessary and sufficient to induce Xenopus tail regeneration. Development *134*, 1323-335.

Adams, D.S., Robinson, K.R., Fukumoto, T., Yuan, S., Albertson, R.C., Yelick, P., Kuo, L., McSweeney, M., and Levin, M. (2006). Early, H+-V-ATPase-dependent proton flux is necessary for consistent left-right patterning of non-mammalian vertebrates. Development *133*, 1657-671.

Agata, K., Nakajima, E., Funayama, N., Shibata, N., Saito, Y., and Umesono, Y. (2006). Two different evolutionary origins of stem cell systems and their molecular basis. Semin Cell Dev Biol *17*, 503-09.

Amigorena, S., Choquet, D., Teillaud, J.L., Korn, H., and Fridman, W.H. (1990). Ion channel blockers inhibit B cell activation at a precise stage of the G1 phase of the cell cycle. Possible involvement of K+ channels. J Immunol *144*, 2038-045.

Apáti, A., Pászty, K., Erdei, Z., Szebényi, K., Homolya, L., and Sarkadi, B. (2012). Calcium signaling in pluripotent stem cells. Mol Cell Endocrinol *353*, 57-67.

Arcangeli, A., and Becchetti, A. (2006). Complex functional interaction between integrin receptors and ion channels. Trends Cell Biol *16*, 631-39.

Arcangeli, A., Crociani, O., Lastraioli, E., Masi, A., Pillozzi, S., and Becchetti, A. (2009). Targeting ion channels in cancer: a novel frontier in antineoplastic therapy. Curr Med Chem *16*, 66-93.

Atkinson, D.L., Stevenson, T.J., Park, E.J., Riedy, M.D., Milash, B., and Odelberg, S.J. (2006). Cellular electroporation induces dedifferentiation in intact newt limbs. Dev Biol *299*, 257-271.

Beane, W.S., Morokuma, J., Adams, D.S., and Levin, M. (2011). A chemical genetics approach reveals H,K-ATPase-mediated membrane voltage is required for planarian head regeneration. Chem Biol *18*, 77-89.

Binggeli, R., and Weinstein, R.C. (1986). Membrane potentials and sodium channels: hypotheses for growth regulation and cancer formation based on changes in sodium channels and gap junctions. J Theor Biol *123*, 377-401.

Blackiston, D.J., McLaughlin, K.A., and Levin, M. (2009). Bioelectric controls of cell proliferation: ion channels, membrane voltage and the cell cycle. Cell Cycle *8*, 3519-528.

Chiu, S.Y., and Wilson, G.F. (1989). The role of potassium channels in Schwann cell proliferation in Wallerian degeneration of explant rabbit sciatic nerves. J Physiol *408*, 199-222.

Cone, C.D. (1969). Electroosmotic interactions accompanying mitosis initation in sarcoma cells in vitro. Trans N Y Acad Sci *31*, 404-427.

Cowles, M.W., Hubert, A., and Zayas, R.M. (2012). A Lissencephaly-1 homologue is essential for mitotic progression in the planarian Schmidtea mediterranea. Dev Dyn *241*, 901-910.

DeCoursey, T.E., Chandy, K.G., Gupta, S., and Cahalan, M.D. (1984). Voltage-gated K+ channels in human T lymphocytes: a role in mitogenesis? Nature *307*, 465-68.

Dimmitt, J., and Marsh, G. (1952). Electrical control of morphogenesis in regenerating Dugesia tigrina. II. Potential gradient vs. current density as control factors. J Cell Physiol *40*, 11-23.

Doyon, J.B., Zeitler, B., Cheng, J., Cheng, A.T., Cherone, J.M., Santiago, Y., Lee, A.H., Vo, T.D., Doyon, Y., et al. (2011). Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells. Nat Cell Biol *13*, 331-37.

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

Fewtrell, C. (1993). Ca2+ oscillations in non-excitable cells. Annu Rev Physiol *55*, 427-454.

Freedman, B.D., Price, M.A., and Deutsch, C.J. (1992). Evidence for voltage modulation of IL-2 production in mitogen-stimulated human peripheral blood lymphocytes. J Immunol *149*, 3784-794.

Guo, T., Peters, A.H., and Newmark, P.A. (2006). A Bruno-like gene is required for stem cell maintenance in planarians. Dev Cell *11*, 159-169.

Hille, B. (2001). Ion channels of excitable membranes (Sunderland, Mass.: Sinauer).

Hübner, C.A., and Jentsch, T.J. (2002). Ion channel diseases. Hum Mol Genet *11*, 2435-445.

Kunzelmann, K. (2005). Ion channels and cancer. J Membr Biol *205*, 159-173.

Ladurner, P., Pfister, D., Seifarth, C., Schärer, L., Mahlknecht, M., Salvenmoser, W., Gerth, R., Marx, F., and Rieger, R. (2005). Production and characterisation of cell- and tissue-specific monoclonal antibodies for the flatworm Macrostomum sp. Histochem Cell Biol *123*, 89-104.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Lang, F., Friedrich, F., Kahn, E., Wöll, E., Hammerer, M., Waldegger, S., Maly, K., and Grunicke, H. (1991). Bradykinin-induced oscillations of cell membrane potential in cells expressing the Ha-ras oncogene. J Biol Chem *266*, 4938-942.

Levin, M. (2009). Bioelectric mechanisms in regeneration: Unique aspects and future perspectives. Semin Cell Dev Biol *20*, 543-556.

Liu, P.Q., Chan, E.M., Cost, G.J., Zhang, L., Wang, J., Miller, J.C., Guschin, D.Y., Reik, A., Holmes, M.C., et al. (2010). Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. Biotechnol Bioeng *106*, 97-105.

Mahon, M.J. (2011). pHluorin2: an enhanced, ratiometric, pH-sensitive green florescent protein. Adv Biosci Biotechnol *2*, 132-37.

Marsh, G., and Beams, H.W. (1947). Electrical control of growth polarity in regenerating Dugesia tigrina. Fed Proc *6*, 163.

Marsh, G., and Beams, H.W. (1952). Electrical control of morphogenesis in regenerating Dugesia tigrina. I. Relation of axial polarity to field strength. J Cell Physiol *39*, 191-213.

Morgan, T.H. (1898). Experimental Studies of the Regeneration of Planaria Maculata. Arch. Entwicklungsmech. Org. *7*, 364-397.

Morgan, T.H. (1902). Growth and regeneration in Planaria lugubris. Arch. Ent. Mech. Org. *13*

*M***outon, S.,** Willems, M., Braeckman, B.P., Egger, B., Ladurner, P., Schärer, L., and Borgonie, G. (2009). The free-living flatworm Macrostomum lignano: a new model organism for ageing research. Exp Gerontol *44*, 243-49.

De Mulder, K., Kuales, G., Pfister, D., Egger, B., Seppi, T., Eichberger, P., Borgonie, G., and Ladurner, P. (2010). Potential of Macrostomum lignano to recover from gamma-ray irradiation. Cell Tissue Res *339*, 527-542.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., et al. (2009). Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. Dev Biol *334*, 198-212.

Nogi, T., and Levin, M. (2005). Characterization of innexin gene expression and functional roles of gap-junctional communication in planarian regeneration. Dev Biol *287*, 314-335.

Nogi, T., Zhang, D., Chan, J.D., and Marchant, J.S. (2009). A novel biological activity of praziquantel requiring voltage-operated Ca2+ channel beta subunits: subversion of flatworm regenerative polarity. PLoS Negl Trop Dis *3*, e464.

Oviedo, N.J., and Levin, M. (2007). smedinx-11 is a planarian stem cell gap junction gene required for regeneration and homeostasis. Development *134*, 3121-131.

Oviedo, N.J., Nicolas, C.L., Adams, D.S., and Levin, M. (2008). Live Imaging of Planarian Membrane Potential Using DiBAC4(3). CSH Protoc *2008*, pdb.prot5055.

Pandiella, A., Magni, M., Lovisolo, D., and Meldolesi, J. (1989). The effect of epidermal growth factor

on membrane potential. Rapid hyperpolarization followed by persistent fluctuations. J Biol Chem *264*, 12914-921.

Paredes-Gamero, E.J., Barbosa, C.M., and Ferreira, A.T. (2012). Calcium signaling as a regulator of hematopoiesis. Front Biosci (Elite Ed) *4*, 1375-384.

Pearson, B.J., and Sánchez Alvarado, A. (2010). A planarian p53 homolog regulates proliferation and self-renewal in adult stem cell lineages. Development *137*, 213-221.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. (2007). The exceptional stem cell system of Macrostomum lignano: screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. Front Zool *4*, 9.

Poss, K.D. (2010). Advances in understanding tissue regenerative capacity and mechanisms in animals. Nat Rev Genet *11*, 710-722.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science *310*, 1327-330.

Sánchez Alvarado, A. (2004). Regeneration and the need for simpler model organisms. Philos Trans R Soc Lond B Biol Sci *359*, 759-763.

Schuster, S., Marhl, M., and Höfer, T. (2002). Modelling of simple and complex calcium oscillations. From single-cell responses to intercellular signalling. Eur J Biochem *269*, 1333-355.

Sheng, M., and Pak, D.T. (2000). Ligand-gated ion channel interactions with cytoskeletal and signaling proteins. Annu Rev Physiol *62*, 755-778.

Sundelacruz, S., Levin, M., and Kaplan, D.L. (2008). Membrane potential controls adipogenic and osteogenic differentiation of mesenchymal stem cells. PLoS One *3*, e3737.

Thomas, A.P., Bird, G.S., Hajnóczky, G., Robb-Gaspers, L.D., and Putney, J.W. (1996). Spatial and temporal aspects of cellular calcium signaling. FASEB J *10*, 1505-517.

Tonelli, F.M., Santos, A.K., Gomes, D.A., da Silva, S.L., Gomes, K.N., Ladeira, L.O., and Resende, R.R. (2012). Stem cells and calcium signaling. Adv Exp Med Biol *740*, 891-916.

Wang, E., Yin, Y., Zhao, M., Forrester, J.V., and McCaig, C.D. (2003). Physiological electric fields control the G1/S phase cell cycle checkpoint to inhibit endothelial cell proliferation. FASEB J *17*, 458-460.

Wang, J., Friedman, G., Doyon, Y., Wang, N.S., Li, C.J., Miller, J.C., Hua, K.L., Yan, J.J., Babiarz, J.E., et al. (2012). Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. Genome Res

Wolff, C., Fuks, B., and Chatelain, P. (2003). Comparative study of membrane potential-sensitive fluorescent probes and their use in ion channel screening assays. J Biomol Screen *8*, 533-543.

Zhang, D., Chan, J.D., Nogi, T., and Marchant, J.S. (2011). Opposing roles of voltage-gated Ca2+ channels in neuronal control of regenerative patterning. J Neurosci *31*, 15983-995.

Zhao, Y., Araki, S., Wu, J., Teramoto, T., Chang, Y.F., Nakano, M., Abdelfattah, A.S., Fujiwara, M., Ishihara, T., et al. (2011). An expanded palette of genetically encoded Ca^{2+} indicators. Science *333*, 1888-891.

CHAPTER 5

WORMS UNDER STRESS: EFFECTS OF LOW SALINITY AND STARVATION ON THE FLATWORM *MACROSTOMUM LIGNANO*

Daniil Simanov[1] and Eugene Berezikov[1,2].

[1] Hubrecht Institute, Hubrecht Institute, KNAW and University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands
[2] European Research Institute for the Biology of Ageing, University of Groningen, University Medical Centre Groningen, 9713AV Groningen, The Netherlands

Work in progress.

ABSTRACT

In order to survive and reproduce, every organism must deploy an array of molecular, physiological, and behavioral responses to avoid or counteract detrimental environmental conditions. One of the most crucial variables for all aquatic organisms is the salinity and, consequently, osmotic power of their environment. Changes in salinity affect many aspects of water organism's life including metabolism, growth, reproduction, regeneration abilities and behavior, yet not much is known about the molecular mechanisms responsible for the salinity stress response in animals, and even less – about the long-term effects of low or high salinity on different organisms. Like other estuarine inhabitants, marine flatworm *Macrostomum lignano* can survive significant salinity fluctuations, rapidly adjusting to changes in osmotic environment. In this study we used RNA-seq technique to analyze transcription dynamics in *M. lignano* in response to extremely low salinity conditions. Being exposed to low salinity, *M. lignano* stop to take food, so the effects of salt stress are always accompanied by the ones caused by starvation. We compared RNA-seq data from the worms at low salinity conditions with the expression profiles of starved worms at normal salinity over a time course. Thus the early effects of starvation on *M. lignano* were also characterized.

INTRODUCTION

Every organism must deal with the changes in its environment. And the more unstable its habitat is, the faster it has to respond. In a lot of cases animals would simply flee stressful and potentially harmful conditions, but when it is not possible, their responses to environmental shifts become more complex, involving changes in gene expression that alter both physiology and behavior. Therefore, to make a full picture of stress response mechanisms in given model organism, we must integrate environmental data, gene expression data, and knowledge of our model system (Ankley et al., 2006). Collecting such data from different organisms enables comparison of stress-response mechanisms in animals from different taxa, occupying a broader range of habitat types (Lettieri, 2006).

One of the environmental factors that have a huge impact on the life of all aquatic organisms is the salinity of water they live in. Today the effects of the stress caused by significant salinity changes are better studied in plants than in animals. Mechanisms of salinity tolerance are described in different plant taxa in a great detail on the cellular (Jacoby et al., 2011) and organismal level, and role of many genes in this process is known (Munns and Tester, 2008; Tuteja, 2007; English and Colmer, 2013). Effects of salinity changes on the growth (Wang et al., 1997) and metabolism (Wang et al., 2012), as well as the molecular pathways responsible for the response to the salt stress (Zhang et al., 2012; Fujii and Zhu, 2012) are studied in many laboratories. The effects of environmental salinity changes on behavior, metabolism, reproduction and regeneration were described in many animal species just as well (Donachy and Watabe, 1986; Shock et al., 2009; Hammen and Lum, 1977; Stueckle et al., 2009; Beck et al., 2003; Ern et al., 2012), however much less is known about the actual molecular mechanisms underlying the recognition of salinity changes and the response to the salt stress (Evans, 2010; Fiol and Kültz, 2007). High tolerance to the salinity changes is described in some salt-marsh inhabitants (Donachy and Watabe, 1986; Shock et al., 2009; Gonzalez, 2012) from the physiological point of view, but the genetic mechanisms and signaling pathways responsible for this adaptation largely remain unclear (Reitzel et al., 2008).

Transcriptome studies can be used to investigate the reaction of the organism to an environmental stressor at molecular level (Snape et al., 2004; Ankley et al., 2006; Poynton and Vulpe, 2009). Recently DNA microarrays and quantitative real time PCR were used in different model systems to link ecotoxicological effects (including the ones caused by exposure to salinity changes) to transcription profiles of a large number of genes (Edge et al., 2005; Plusquin et al., 2012). The expression data obtained help to identify the complex pathways altered in an exposed organism, which could lead to the discovery of potential stress and toxicity biomarkers (Steinberg et al., 2008).

Flatworm species are known for their adult pluripotent stem cell system and regeneration capacity. Traditionally, they are used as models in studies concerning stem cell functioning and regeneration (Rink, 2013; Aboobaker, 2011; Egger et al., 2006). At the same time they can be successfully used for drug screening (Buttarelli et al., 2008; Rawls et al., 2010) and eco-toxicology studies (Indeherberg et al., 1999; de Lucca C Preza and Smith, 2001; Navarro et al., 2009). There are many advantages of flatworms as models for studies of the mechanisms of toxic stress reaction, including their ease of culture in the laboratory and sensitivity to low

concentrations of environmental toxins (Nano et al., 2002). Interestingly, previous research has demonstrated that their cellular and physiological responses are similar to those in mammals (Hall et al., 1986; Schaeffer, 1993; Buttarelli et al., 2008), suggesting the same molecular mechanisms being involved.

Macrostomum lignano is an emerging flatworm model that can be used to investigate molecular mechanisms underlying stress tolerance and stress response. It is a small animal (adults are about 1,5 mm long), with simple body plan and short generation time (Ladurner et al., 2005). Worms are also highly transparent, so all the phenotypic changes in response to environmental conditions can be seen through the light microscope *in vivo*. Genome and transcriptome of *M. lignano* are sequenced and publicly available (see www.macgenome.org), so it is possible to study gene expression changes in response to environmental signals (such as salinity fluctuations) genome-wide. Importantly, *M. lignano* is a free-living worm, and its natural habitat is estuarine intertidal areas in the Northern Adriatic (Ladurner et al., 2000) that are characterized by variable salinity regimes due to seasonal and daily fluctuations. *M. lignano* is well adapted to such conditions and can withstand significant salinity changes, which makes this worm an extremely interesting model for salinity tolerance and salt stress studies. Normally worms are cultured in the laboratory in f/2 medium (sea water based) with the salinity of 32‰, but can survive in medium with 1,5‰ of salt. The most obvious and very intriguing effect of such low salinity conditions in *M. lignano* is rapid degradation of the gonads. At the same time diatom algae *Nitzschia curvilineata* that are used as food for the worms cannot survive in low salinity conditions and die. *M. lignano* do not feed on dead algae, so reduction of medium salinity leads to starvation. Starved animals normally also lose their gonads (Nimeth et al., 2004; Pfister et al., 2008), so it is unclear whether gonads degradation is indeed caused by salinity conditions or simply by starvation.

In this study we followed *M. lignano* populations exposed to low salinity in comparison with worms, cultured at normal salinity in the absence of food over a time course. We collected and compared RNA-seq data from both groups of animals at different time points in order to identify changes in gene expression that are caused by salinity change, excluding the effects of starvation. In the Chapter 4 we discussed ion channels in *M. lignano* and their role in stem cell functioning (Simanov et al., 2012). We expect expression of ion channel genes to be altered following osmotic shock and low salinity stress. And we ant to check, whether it would be the same transcripts that are possibly enriched in neoblasts (See Chapter 3) or completely separate set of genes.

Degradation of the gonads and general body shrinkage during starvation was first described in flatworms over a hundred years ago (Morgan, 1902). This process is completely reversible, and when food becomes available again (or after regeneration is finished), the worms normally grow back to their original size, rebuilding their gonads as well (Romero and Baguñà, 1991). Flatworms can also lose their reproductive organs during regeneration, or (some species) during a shift from sexual to asexual reproduction (Hyman, 1951; Fedecka-Bruner, 1967). In all these cases gonads degradation happens as a result of change in cell number, rather than a cell size (Baguñà and Romero, 1981; Oviedo et al., 2003). Furthermore, there is no decrease in the number of proliferating stem cells in response to starvation (Baguñà, 1976; Orii et al.,

2005), meaning that cell death must explain the shrinkage of planarians during starvation. Two different cell death mechanisms might be responsible for that. Cases of autophagy (lysis of cells followed by degradation of whole organelles) in planarian flatworms can be observed as early as the first week of starvation (Bowen and Ryder, 1974; Bowen et al., 1982; Bowen et al., 1976), and the clear increase in number of apoptotic cells is detected after 5 weeks of starvation (Pellettieri et al., 2010). Functional relationship between these cell death processes (autophagy and apoptosis) is very complex (Levine and Yuan, 2005; Baehrecke, 2005), and the exact mechanism responsible for the gonads degradation during starvation in flatworms still remains to be investigated. Thus, another goal of this study is to use RNA-seq data obtained from starved worms to get an insight into the early activation and regulation of mechanisms involved in gonad degradation and body shrinkage.

RESULTS AND DISCUSSION

1. Effect of the low salinity on *M. lignano*

To study the effects of extremely low salinity on *M. lignano*, we put a population of worms into f/2 medium diluted 20 times with fresh water, reaching the salinity of 1,5 – 1,8‰. At these conditions around 6% of worms die within first three hours of being exposed to low salinity (data not shown), while the rest of the worms survive at least for 2 weeks (the worms were never followed longer). At the salinity of 1‰ all the worms die within first 3 hours, suggesting that 1,5–1,8‰, used in this study, is the extreme salinity on the border of *M. lignano* tolerance. During the first hour of being exposed to low salinity worms move much slower than normally or do not move at all. They also look more roundish than control animals (worms put to the fresh plates with no food at the same moment when the other group was exposed to the low salinity). Because of the round body shape it is hard to see and estimate the size of the gonads under the light microscope (Fig. 1). After 2-3 hours worms seem to recover from the initial shock, and 15 hours after the start of the experiment they look and behave normally, and there is no visible difference between them and control animals (Fig. 1). However, after 60 hours the gonads of the worms at low salinity conditions seem to be almost completely degraded, while the starved worms still look intact and even lay eggs (Fig. 1). 96 hours after the start of the experiment, no gonads can be identified in worms kept in low salinity medium, and starved worms still have their ovaries and testis for another 5 to 7 days (data not shown).

Thus, gonad degradation in worms exposed to the low salinity appears to develop much faster than in simple case of starvation. Stress caused by low salinity results in the activation of mechanisms other than the ones involved in gonad degradation in the course of starvation, or at least the initial triggering step of this mechanism is different.

2. Characterization of time points for RNA-seq data sets

To perform a whole transcriptome RNA-seq we extracted total RNA from the worm populations, both starved and exposed to low salinity conditions, at three different time points, mentioned above. At 1 hour after being moved to the low salinity medium worms appear to be in shock. They move slowly or do not move at all, and their body shape is slightly changed. At this time

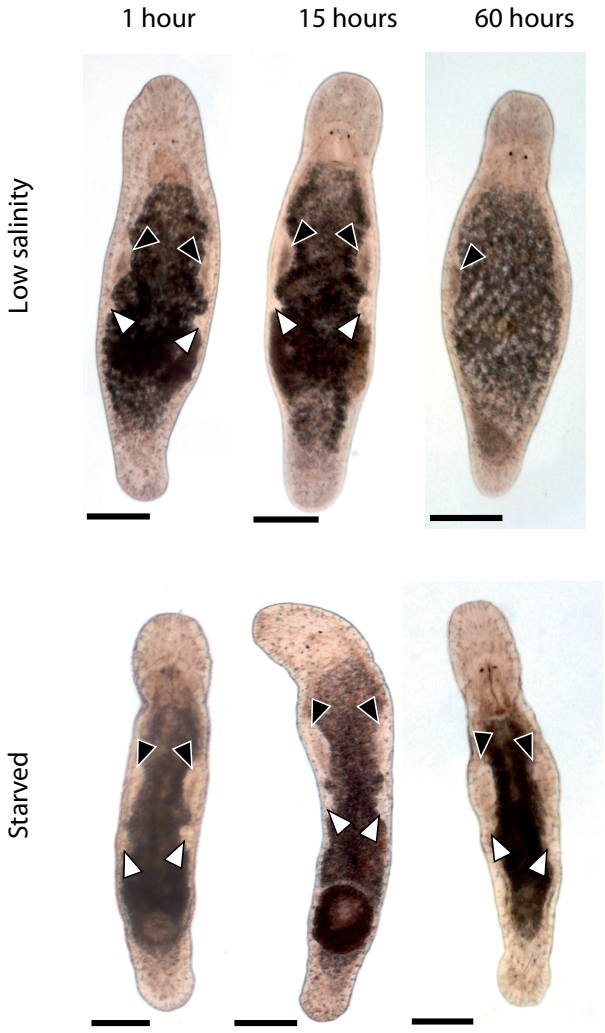


Figure 1. Short-term effects of low salinity and starvation on *M. lignano*. Worms at low salinity conditions start losing gonads earlier than starved worms. At 60 hours after the start of the experiment most of the animals in low salinity medium have no ovaries. More explanation in the text. Black arrowheads indicate testis, white arrowheads indicate ovaries. Scalebar 200 µm.

point we expect salt stress response genes expression to be induced. However, *M. lignano* adjust to this extreme salinity change very quickly (see above), suggesting that at least some stress-related mechanisms are induced at the posttranslational level, and it is rather proteins than genes that get activated. Still it is important to identify early effects of salinity stress at the gene expression level and find molecular mechanisms that get activated or repressed immediately in response to the salinity change. This might also provide an insight into the signaling mechanisms responsible for sensing the osmotic changes of the medium and

translating them into gene expression programs. At 15 hours after start of the experiment the "shock phenotype" is completely gone, and worms look almost the same as control animals. At this time point most of the stress-related effects should be gone, and it is expression of genes responsible for executing salinity tolerance and re-adjusting worm metabolism to low salinity conditions that is likely to be changed. The size of the gonads at this moment is not yet changed in comparison with controls, but some mechanisms responsible for later gonad degradation might already get activated as well. At 60 hours after initial exposure to the low salinity conditions gonads of *M. lignano* are obviously reduced in size and in some cases are not visible anymore. Worms look and behave normally otherwise. By this time we expect most of the metabolic processes to be adjusted to low salinity conditions. Mechanisms responsible for gonad degradation should be active as well.

1 hour after the start of the experiment starved worms still have their guts full of algae, and digestion is still in process. There is virtually no difference between these animals and worms that still have access to food. Thus, we use the transcriptome of these animals as ultimate control and reference point for all other RNA-seq data sets. 15 hours after the start of starvation most of the worms have no algae left in their guts, yet it is still too early for starvation effects to kick in on the gene expression level. 60 hours after the animals were removed from food they should start adapting their metabolism (and, probably, behavior) for starvation conditions. Some early genes responsible for autophagy processes might also get activated at this time point (Bowen et al., 1982). We expect no apoptosis-related programs to be active yet (since in planarians increase in apoptotic cell number is not observed until the 5th week of starvation) (Pellettieri et al., 2010). Gonads are not reduced in size at this moment, and are even still functioning, so no genes responsible for gonad shrinkage should be activated either.

By comparing RNA-seq data sets from starved worms and animals introduced to low salinity conditions at three early time points we can get an insight into specific genetic mechanisms responsible for a few intriguing physiological processes. Genes expressed differently in the worms just put into the low salinity medium in comparison with worms just removed from food are likely to be involved mainly in stress-response programs. 15 hours after the start of the experiment expression of these genes should get back to normal levels (thus, different from "low salinity, 1 hour" and close to "starved, 1 hour" conditions), while genetic program responsible for high salinity tolerance in *M. lignano* should become active (expression of these genes we expect to be up- or down-regulated in comparison with both starved and "low salinity" worms at the 1st hour of the experiment). Same programs are likely to be still active 60 hours after the animals were moved to low salinity conditions, and genes responsible for gonad degradation should already be active as well. Starvation already has its effects on these worms, but we can identify starvation-related genes (and exclude them from our search for low salinity-specific ones) by comparison with expression data from starved worms at the same time point. We do not expect significant changes in gene expression in starved worms between 1 and 15 hours after the start of the experiment (since some worms even still have food in their guts at this point), while set of genes expressed differently at 60 hours of starvation (in comparison with the 1st hour) is likely to be involved in starvation-related metabolic changes and, probably, early autophagy-activation programs.

3. RNA-seq libraries

12 RNA-seq libraries were constructed and sequenced (two replicas for every time point of both starvation and low salinity conditions).

Between 9,7 mln and 54,1 mln high-quality reads were obtained for each library (Table 1), of which 52%-83% were mappable to the *de novo* transcriptome assembly MLRNA131024 (see Chapter 3). Transcript abundance was inferred from read mappings using RSEM software (Li and Dewey, 2011) and expressed in transcripts per million, or TPM. This metric allows comparison of gene expression levels both within one sample and between different samples. The distribution of gene expression levels is similar in all libraries, with the only deviation in sample 'Starved_15h_4' (Fig. 2A). Hierarchical clustering of the samples (Fig. 2B) shows the expected separating between starved and low salinity samples, and replicates cluster together with the exception of the same 'Starved_15h_4' library. Although this library appears to be biased, we still included it in the subsequent analysis of differential gene expression, reasoning that the bias is not extreme.

4. Comparisons of RNA-seq data sets

We compared expression data from the animals put in low salinity or starvation conditions at different time points. We looked at the number of genes that were expressed differently in different situations, and also at the function of these genes according to GO terms where

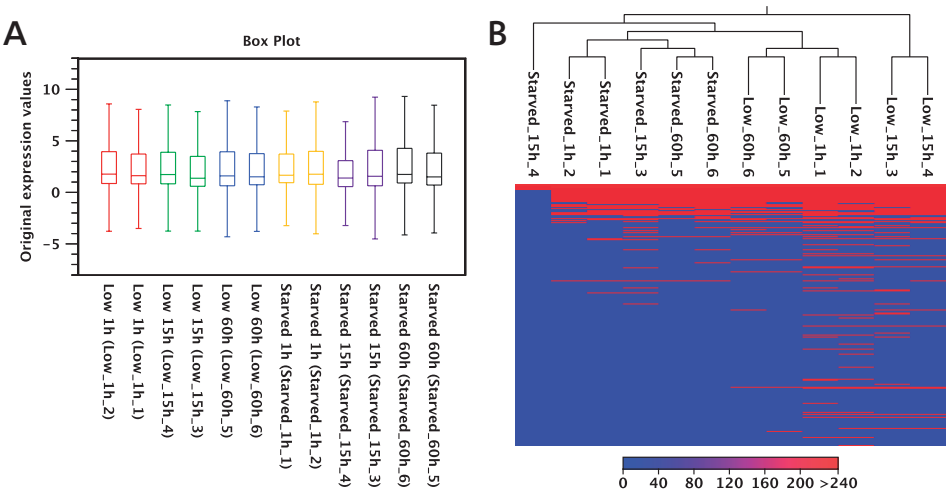


Figure 2. General characteristics of RNA-seq libraries. (A) Box plot showing distribution of expression values of the libraries. No significant deviations between the libraries observed with the exception of 'Starved_15h_4', which shows smaller distribution range. **(B)** Heatmap of the 12 libraries based on 20483 transcripts differentially expressed between at least one pair of conditions. Hierarchical clustering of the samples was performed with using Euclidian distance and average linkage algorithms. Libraries from starved worms group together and libraries from "low salinity" worms as well, regardless of the time points. For all cases but one, replica libraries (from the same time point and same conditions) group together nicely. The only exception is the 'Starved_15h_4', which is also the outlier in the Box plot analysis.

Table 1. Sequencing depth and fraction of mapped reads in RNA-seq libraries.

Library	Preprocessed reads	Mapped reads	% Mapped
Low_1h_1	43,088,592	35,209,653	81.71
Low_1h_2	17,005,184	12,368,225	72.73
Low_15h_3	32,536,828	26,887,516	82.64
Low_15h_4	12,871,625	9,141,754	71.02
Low_60h_5	12,743,835	6,670,870	52.35
Low_60h_6	54,106,830	38,897,900	71.89
Starved_1h_1	25,111,776	20,606,041	82.06
Starved_1h_2	10,338,725	5,570,484	53.88
Starved_15h_3	23,934,498	19,958,358	83.39
Starved_15h_4	9,716,456	5,465,778	56.25
Starved_60h_5	42,303,729	35,155,606	83.10
Starved_60h_6	58,705,764	43,186,974	73.57

identified human homologs are available. The MLRNA131024 *M. lignano* transcriptome assembly contains in total 174,922 transcripts, including gene isoforms. Of these, expression of 170,022 transcripts is detected in current RNA-seq dataset, and 112,354 transcripts have expression value of at least 1 TPM in at least one of the samples. In total, there are 51,625 transcripts in the transcriptome assembly with identifiable human homologs (representing 10,269 non-redundant human genes), and of these 11,057 transcripts (or roughly 1/5th of transcripts with human homologues) are differentially expressed between at least one pair of studied conditions. While transcripts with human homologs represent only a fraction of all genes differentially expressed in this study, this fraction is very significant and in the absence of better alternatives information about human homologs is a good proxy for obtaining an insight about biological processes and functions affected by the experimental conditions.

4.1. Early changes in gene expression. Already 1 hour after the start of the experiment there is a difference in expression of a vast set of genes between worms put into low salinity medium and starved animals (Fig. 3A). Expression of 3475 transcripts has changed more than 2-fold (PPDE >= 0.95), with 2066 transcripts up-regulated in low salinity and 1409 transcripts down-regulated in low salinity compared to control. Of these, 715 and 503 transcripts respectively have identifiable human homologues, and the analysis of GO-term enrichment using human gene annotations shows that many of these genes are involved in transmembrane ion transport (Sup. Fig. 1 and 2, Sup. Tables 1 and 2). However there is also a large group of genes involved in signal transduction during stress response reactions, including mostly receptors (GO:0007165; GO:0007166, GO:0007169). Interestingly, one of the processes affected already at this stage is stress-related apoptosis, and expression of a few caspases (such as caspases 3, 9 and 10) genes, as well as other cell death-associated genes is changed significantly (GO:0097190 and GO:2001236 in Sup. Table 2). Thus, it is both stress-related and low-salinity tolerance responsible genes that are likely to be affected as early as at the 1st hour after the

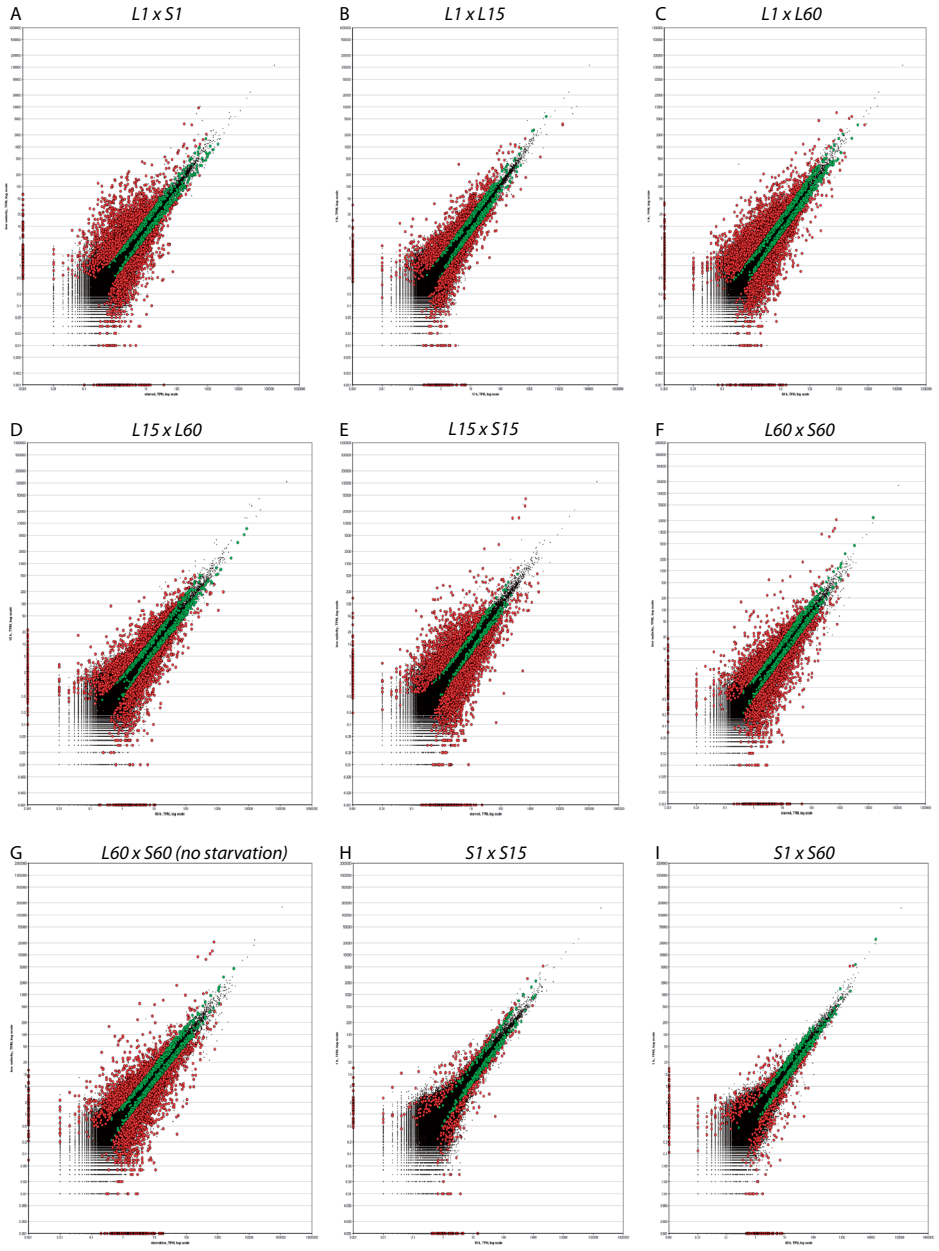


Figure 3. Distribution of expression values between RNA-seq data sets from worms at different conditions and different time points. Each graph shows two RNA-seq data sets compared. Each dot represents a transcript, and the axis show expression values in transcripts per million (TPM, logarithmic scale) in two data sets being compared. Genes with posterior probability of differential expression of at least 0.95 are show in red (more than 2-fold differences) and in green (less than 2-fold difference). **(A)** Comparison of RNA-seq data from the worms at low salinity (y-axis) and starved worms (x-axis), both at 1 hour after the start of the experiment. **(B)** Worms at low salinity, 1 hour (y-axis) and 15 hours (x-axis) after the medium was changed. **(C)** Low salinity, ▶

change of the medium, and it might be that certain genes involved in gonad degradation later, are also activated already at this stage.

To further address the question of early response genes, we identified a subset of 1331 transcripts that are differentially expressed between low salinity and starvation at 1 hour but are not differentially expressed at later stages (15h and 60h). GO term analysis of these transcripts revealed enrichment in biological processes 'cellular response to cAMP', 'protein dephosphorylation' and 'interferon signaling pathway' among upregulated genes (Sup. Fig. 8; Sup. Table 8), whereas among down-regulated transcripts there is enrichment for ion transport activity (Sup. Fig. 7; Sup. Table 7). This subset of genes appears the most likely to be involved in early adaptation to low salinity rather than in the maintenance of this adaptation. Remarkably, there is an obvious change in expression of genes involved in cell adhesion and movement, most likely caused by mechanical stress and damage (Sup. Fig. 7; Sup. Table 7).

4.2. Changes in gene expression during adaptation to low salinity. Expression of a large set of genes (2265 transcripts changed 2-fold or more) is changed at low salinity conditions between 1 and 15 hours after the start of the experiment (Fig. 3B, Fig. 4), and another big change happens between 15 and 60 hours with 3060 transcripts changing their expression 2-fold or more (Fig. 3D, Fig. 4). When comparing expression profiles of worms at low salinity 1 hour and 60 hours after the medium was changed, 5862 transcripts appear to change their expression more than 2-fold (Fig. 3C, Fig. 4), which is more than ones changed between 1 hour and 15 hours and between 15 hours and 60 hours taken together. Changes in expression profiles are accumulating during first 60 hours in low salinity medium. Genes changing their expression belong mainly to one of the groups mentioned above - the ones responsible for ion transport and metabolism. At 15 and 60 hours time points, as expected, no genes responsible for stress response are found anymore among the ones that are up- or down-regulated (data not shown).

4.3. Gene expression changes during starvation. Much less changes in expression occur during first 60 hours of starvation. Expression of 666 transcripts is changed 2-fold or more between 1h and 60h (Fig 3I, Fig. 4). As expected, mainly these are the genes involved in metabolism (for example, lipid, amino-acid or nucleotide metabolism), but there are also signaling pathways affected by starvation. For example, Notch gets down-regulated (Sup. Fig. 3 and 4; Sup. Tables 3 and 4). There are virtually no apoptosis-related genes among the ones with expression significantly changed. However, there are a few early endosome and lysosome markers (such as filamin alpha and Rab7) on the list of upregulated genes, and it might be an indication of first autophagy events (GO:0045022, Sup. Table 3).

▶ 1 hour (y-axis), versus low salinity, 60 hours (x-axis). **(D)** Low salinity, 15 hours (y-axis), and low salinity, 60 hours (x-axis). **(E)** Low salinity, 15 hours (y-axis), and starved, 15 hours (x-axis). **(F)** Low salinity, 60 hours (y-axis), and starved, 60 hours (x-axis). **(G)** Same as (F), but the starvation-related genes were taken out of the comparison. **(H)** Starved worms, 1 hour (y-axis), and starved, 15 hours (x-axis). **(I)** Starved, 1 hour (y-axis), and starved, 60 hours (x-axis).

4.4. Gene expression changes in long-term salinity adaptation. A large group of genes is expressed in worms at low salinity conditions differently than in starved worms at 60 hours after the start of the experiment. Expression of 3243 transcripts is changed 2-fold or more (Fig. 3F, Fig. 4). Most of these genes are involved in ion transport (Sup. Fig. 5 and 6; Sup. Tables 5 and 6) and overlap largely with the ones that we saw among up- or down-regulated in "low salinity" worms at 60 hours in comparison with the first hour of the experiment. We think, these are mostly the genes responsible for outstanding salinity tolerance of *M. lignano*. There is also a large group of genes involved in metabolic processes, which is also not surprising. Even though both groups of animals are technically starved (since they do not have access to food), their general metabolic rates might differ significantly under the influence of salinity change. Active transport of solutes against emerging osmotic gradient is energetically expensive, and low salinity conditions were shown to result in an increase in metabolic rate (Haney et al., 1999). There is a small group of apoptosis-related genes such as caspases 8 and 10 that are also expressed differently in "low salinity" and starved worms. Importantly, among the

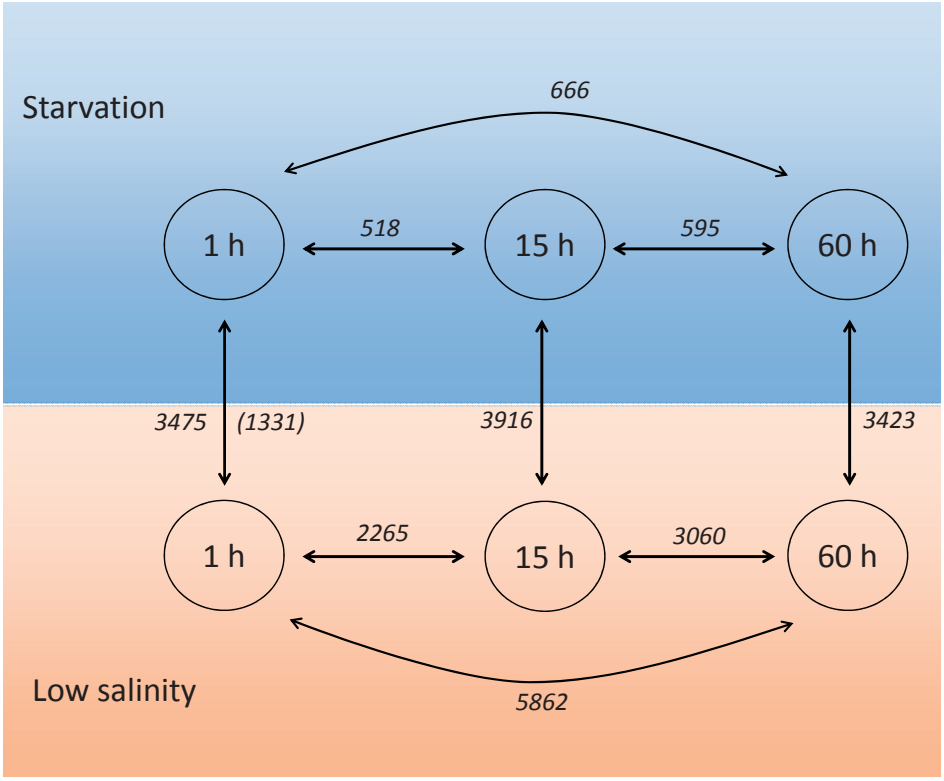


Figure 4. Scheme of the comparative analysis. Arrows with numbers represent RNA-seq data comparisons with the number of genes up- or down-regulated at least 2-fold. Number in brackets shows the genes specifically involved in the early response to low salinity. Their expression is changed at least 2-fold at the first hour on low salinity in comparison with the control, but is not affected at later time points. More information in the text.

genes significantly down-regulated at low salinity conditions there are more than 30 involved in gonad formation and functioning (GO:0022414 - CDKL2, HSPA2, SPA17, Sup. Table 5). It might be that these very genes play role in the early degradation of gonads. Finally, some transcripts encoding cell contact proteins and proteins involved in cytoskeleton formation are up-regulated at low salinity conditions (Sup. Table 6).

CONCLUSION AND FUTURE DIRECTIONS

In this work we described a previously unknown effect of low salinity conditions on the reproductive system of the flatworm *M. lignano*, showing that in low salinity medium worms rapidly lose their gonads - much faster than in the situation of food absence. In order to study this and other effects of low salinity stress on the molecular level we generated RNA-seq data sets from the worms stressed by starvation or low salinity conditions at different time points within first 3 days after conditions switch.

As was expected, low salinity conditions trigger massive changes in gene expression transcriptome-wide. At 1 hour after the start of the experiment hundreds of transcripts involved in metabolism and ion-transport are up- or down-regulated already. We consider these genes (and the ones activated later) to be responsible for impressive salinity tolerance in *M. lignano*. There are a few of them that are not expressed under normal conditions, and get immediately activated in low salinity, and the ones that, on the contrary, get silenced completely once the medium is changed (Fig 3A, red dots on the axis themselves). These genes, working only at certain salinity and activated or silenced very quickly might be extremely important for execution of salinity tolerance program. There are not many of them and further studies of these genes (finding out expression patterns and functions) would help to understand mechanisms of low salinity tolerance better. It is also important to add a few later time points to the experiment and generate more RNA-seq data to see when the gene expression stabilizes and to distinguish transcripts involved in establishing salinity tolerance and the ones maintaining it.

As expected, expression of a large set of ion channel genes is changed at every time point after the start of the experiment. Thus, ion channels are involved in osmotic stress reaction, adaptation to low salinity and exceptional salinity tolerance of *M. lignano*. We plan to compare the set of ion channel genes involved in low salinity response with genes, potentially enriched in neoblasts and involved in stem cell functioning (see Chapters 3 and 4). It would be interesting to see whether ion channels actively expressed in stem cells could also play role in basic osmotic regulation, or their functions are specific for neoblasts.

All the experiments described should also be repeated with high salinity stress conditions, for neither tolerance nor the effect of high salinity on *M. lignano* was ever studied. It would be interesting to see whether there are the same mechanisms responsible for low and high salinity stress responses.

Within first 3 days of starvation worms do not change phenotypically, maintaining same body size and proportions as normal well-fed animals. Accordingly, there are not many changes happening on gene expression level. However, some changes in metabolism are induced at the very early stages of starvation, and first possible signs of autophagy induction can be seen as early as 60 hours after the start of starvation. In accordance with the results obtained from

planarian flatworms (Pellettieri et al., 2010), no signs of apoptosis can be detected at the early stages of starvation, even at gene expression level.

We identified a group of genes that is potentially involved in early gonad degradation in worms at low salinity conditions. These 32 transcripts encoding proteins involved in reproductive system functioning are worth looking at in more detail. It would be interesting to see whether the same set of genes is involved in re-establishing the gonads once the worms are put back to normal salinity conditions. Finally, it still remains unclear which genes play a role of the switch, triggering gonad degradation. One good candidate here would be Notch signaling pathway since the expression of several Notch and Notch receptor isoforms is significantly changed at different time points after the worms are put into low salinity medium. Even if Notch signaling does not affect the gonad degradation it would be still interesting to understand the role of this pathway in low salinity response.

MATERIALS AND METHODS

1. Animal culture

M. lignano worms of inbred DV1 line (Janicke et al., 2013) were cultured in Petri dishes in f/2 medium, nutrient-enriched artificial sea water (Andersen et al., 2005), 150-200 animals per plate. All the worms were young adults, 3-5 weeks after hatching. Worms were fed with unicellular diatom *Nitzschia curvilineata*, which covers the bottom of the petri dishes (Ladurner et al., 2005). All animals were kept in the incubators under following conditions: 20°C temperature, 60% humidity and 14/10 h day/night cycle (Rieger et al., 1988). For starvation experiments worms were simply moved to clean plates with no algae and fresh f/2 medium. In low salinity experiments animals were first moved to clean plates with no food and a small volume of fresh f/2. Then f/2 medium was gradually, within 1 hour, diluted with MiliQ water to the salinity of 1,5 - 1,8‰. Starved worms and worms at low salinity were kept in the same incubator, RNA isolations (at all 3 time points) were performed on them simultaneously. First time point (1 hour) for the starved worms was exactly 1 hour after the worms were moved to clean plates; for the "low salinity" worms time was counted from the moment f/2 medium started to be diluted, so 1 hour time point is technically the moment when salinity of the medium reached its minimum of 1,5 - 1,8‰. Replicate experiments were performed one month after the initial ones, on independent batch of worms.

2. Total RNA extraction

Total RNA was extracted using TRI Reagent (T9424, Sigma), according to manufacturer's instructions. Animals were homogenized in TRI Reagent by pipetting. For every extraction a batch of 150-200 worms was used (all the worms from the plate). Samples were resuspended in nuclease-free water and treated with 5 U of DNAse I (Thermo Scientific, Cat. No. EN0521) for 45 minutes at 37°C. Enzyme and all the remaining DNA were removed by extraction with phenol : chloroform : isoamyl alchohol (125:24:1, pH 4,5 Life technologies, Cat. No. AM9720). Samples were alcohol precipitated overnight at −80°C. Total RNA was pelleted by centrifugation at 12,000g for 20 mins at 4°C, washed with 70% ethanol and air-dried for 5 minutes. RNA was

resuspended in nuclease-free water. Concentration of total RNA samples was measured with Qubit RNA BR assay kit (Invitrogen, Cat. No. Q10211).

3. Preparation and sequencing of RNA-seq libraries

Poly(A)-tailed mRNA fraction was purified from total RNA and barcoded RNA-seq libraries were created using SureSelect Strand Specific RNA Reagent Kit (Agilent, G9691A) in accordance with manufacturer's protocol. Replicate libraries were made and sequenced independently from the first set of libraries. Sequencing was performed on Illumina HiSeq2500 machine in the RapidRun mode using single end 100 nt protocol. PhiX control libraries were added to every run according to manufacturer's protocol.

4. Data analysis

Raw reads were first preprocessed to filter occasional adapter and phiX sequences using preprocessing module of Mira assembler (Chevreux et al., 2004) and low-quality bases were further clipped using the preprocessing module of SGA assembler (Simpson and Durbin, 2012) with the following parameters: 'sga preprocess --quality-trim=15 --quality-filter=3 --min-len=64'. Preprocessed reads were mapped to MLRNA131024 transcriptome assembly using bowtie v.1.0.0 (Langmead 2009) with the following parameters: '-q -n 2 -e 99999999 -l 16 -a -m 200 --best –strata'. The mapping results were filtered to exclude hits with more than 5 mismatches (NM:I tag in SAM output) and transcript quantification was performed by RSEM package v.1.2.7 (Li and Dewey, 2011), which takes into account read mapping ambiguity when estimating gene expression values. The resulting transcript abundance estimates, expressed in transcripts per million (TMP), were used in calculation of differentially expressed genes by EBSeq software (Leng et al., 2013) with two replicates for each experimental condition and cutoff for posterior probability of differential expression (PPDE) of at least 0.95 (equivalent to false discovery rate of 0.05). Boxplots and heatmaps were generated by CLC Workbench software (www.clcbio.com). GO term enrichment was analyzed using GOrilla service (Eden et al., 2009) in the unranked gene list mode and human homolog gene annotations.

REFERENCES

Aboobaker, A.A. (2011). Planarian stem cells: a simple paradigm for regeneration. Trends Cell Biol *21*, 304-311.

Andersen, R.A., Berges, R.A., Harrison, P.J., and Watanabe, M.M. (2005). Recipes for freshwater and seawater media; enriched natural seawater media. In: Anderson, R.A. (Ed.), Algal culturing techniques, Elsevier academic press, Burlington, 507.

Ankley, G.T., Daston, G.P., Degitz, S.J., Denslow, N.D., Hoke, R.A., Kennedy, S.W., Miracle, A.L., Perkins, E.J., Snape, J., et al. (2006). Toxicogenomics in regulatory ecotoxicology. Environ Sci Technol *40*, 4055-065.

Baehrecke, E.H. (2005). Autophagy: dual roles in life and death? Nature Reviews Molecular Cell Biology *6*, 505-510.

Baguñà, J. (1976). Mitosis in the intact and regenerating planarian Dugesia mediterranea n. sp. I. Mitotic studies during growth, feeding and starvation. Journal of Experimental Zoology *195*, 53-64.

Baguñà, J., and Romero, R. (1981). Quantitative analysis of cell types during growth, degrowth and regeneration in the Planarians Dugesia Mediterranea and Dugesia Tigrina. In The Biology of the Turbellaria, Springer.

Beck, C., Blumer, L., and Brown, T. (2003). Effects of salinity on metabolic rate in black mollies.

Bowen, E.D., Ryder, T.A., and Dark, C. (1976). The effects of starvation on the planarian worm Polycelis tenuis Iijima. Cell Tissue Res *169*, 193-209.

Bowen, I.D., and Ryder, T.A. (1974). Cell autolysis and deletion in the planarian Polycelis tenuis Iijima. Cell Tissue Res *154*, 265-274.

Bowen, I.D., den Hollander, J.E., and Lewis, G.H. (1982). Cell death and acid phosphatase activity in the regenerating planarian Polycelis tenuis Iijima. Differentiation *21*, 160-67.

Buttarelli, F.R., Pellicano, C., and Pontieri, F.E. (2008). Neuropharmacology and behavior in planarians: translations to mammals. Comp Biochem Physiol C Toxicol Pharmacol *147*, 399-408.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E., Wetter, T., and Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res *14*, 1147-159.

Donachy, J.E., and Watabe, N. (1986). Effects of salinity and calcium concentration on arm regeneration byOphiothrix angulata (Echinodermata: Ophiuroidea). Marine Biology *91*, 253-57.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Edge, S.E., Morgan, M.B., Gleason, D.F., and Snell, T.W. (2005). Development of a coral cDNA array to examine gene expression profiles in Montastraea faveolata exposed to environmental stress. Mar Pollut Bull *51*, 507-523.

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

English, J.P., and Colmer, T.D. (2013). Tolerance of extreme salinity in two stem-succulent halophytes (Tecticornia species). Functional Plant Biology

Ern, R., Huong, D.T.T., Nguyen, V.C., Wang, T., and Bayley, M. (2012). Effects of salinity on standard metabolic rate and critical oxygen tension in the giant freshwater prawn (Macrobrachium rosenbergii). Aquaculture Research

Evans, T.G. (2010). Co-ordination of osmotic stress responses through osmosensing and signal transduction events in fishes. J Fish Biol *76*, 1903-925.

Fedecka-Bruner, B. (1967). [Differentiation of the male gonads in the planarian, Dugesia lugubris, during regeneration]. C R Seances Soc Biol Fil *161*, 21-23.

Fiol, D.F., and Kültz, D. (2007). Osmotic stress sensing and signaling in fishes. FEBS J *274*, 5790-98.

Fujii, H., and Zhu, J.K. (2012). Osmotic stress signaling via protein kinases. Cell Mol Life Sci *69*, 3165-173.

Gonzalez, R.J. (2012). The physiology of hyper-salinity tolerance in teleost fish: a review. J Comp Physiol B *182*, 321-29.

Hall, F., Morita, M., and Best, J.B. (1986). Neoplastic transformation in the planarian: I. Cocarcinogenesis and histopathology. J Exp Zool *240*, 211-227.

Hammen, C.S., and Lum, S.C. (1977). Salinity tolerance and pedicle regeneration of Lingula. Journal of Paleontology , 548-551.

Haney, D.C., Nordlie, F.G., and Binello, J. (1999). Influence of simulated tidal changes in ambient salinity on routine metabolic rate in Cyprinodon variegatus. Copeia , 509-514.

Hyman, L.H. (1951). The Invertebrates: Platyhelminthes and Rhynchocoela: The Acoelomate Bilateria (McGraw-Hill New York).

Indeherberg, M.B., van Straalen, N.M., and Schockaert, E.R. (1999). Combining life-history and toxicokinetic parameters to interpret differences in sensitivity to cadmium between populations of Polycelis tenuis (Platyhelminthes). Ecotoxicol Environ Saf *44*, 1-11.

Jacoby, R.P., Taylor, N.L., and Millar, A.H. (2011). The role of mitochondrial respiration in salinity tolerance. Trends Plant Sci *16*, 614-623.

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D.B., and Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution *67*, 3233-242.

Ladurner, P., Rieger, R., and Baguñà, J. (2000). Spatial distribution and differentiation potential of stem cells in hatchlings and adults in the marine platyhelminth macrostomum sp.: a bromodeoxyuridine analysis. Dev Biol *226*, 231-241.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics *29*, 1035-043.

Lettieri, T. (2006). Recent applications of DNA microarray technology to toxicology and ecotoxicology. Environ Health Perspect *114*, 4-9.

Levine, B., and Yuan, J. (2005). Autophagy in cell death: an innocent convict? J Clin Invest *115*, 2679-688.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or

without a reference genome. BMC Bioinformatics *12*, 323.

de Lucca C Preza, D., and Smith, D.H. (2001). Use of newborn Girardia tigrina (Girard, 1850) in acute toxicity tests. Ecotoxicol Environ Saf *50*, 1-3.

Morgan, T.H. (1902). Growth and regeneration in Planaria lugubris. Arch. Ent. Mech. Org. *13*

*M***unns, R.,** and Tester, M. (2008). Mechanisms of salinity tolerance. Annu Rev Plant Biol *59*, 651-681.

Nano, G.M., Binello, A., Bianco, M.A., Ugazio, G., and Burdino, S. (2002). In vitro tests to evaluate potential biological activity in natural substances. Fitoterapia *73*, 140-46.

Navarro, B.S., Michiels, N.K., Köhler, H.-R., and DSouza, T.G. (2009). Differential expression of heat shock protein 70 in relation to stress type in the flatworm Schmidtea polychroa. Hydrobiologia *636*, 393-400.

Nimeth, K.T., Mahlknecht, M., Mezzanato, A., Peter, R., Rieger, R., and Ladurner, P. (2004). Stem cell dynamics during growth, feeding, and starvation in the basal flatworm Macrostomum sp. (Platyhelminthes). Dev Dyn *230*, 91-99.

Orii, H., Sakurai, T., and Watanabe, K. (2005). Distribution of the stem cells (neoblasts) in the planarian Dugesia japonica. Dev Genes Evol *215*, 143-157.

Oviedo, N.J., Newmark, P.A., and Sánchez Alvarado, A. (2003). Allometric scaling and proportion regulation in the freshwater planarian Schmidtea mediterranea. Dev Dyn *226*, 326-333.

Pellettieri, J., Fitzgerald, P., Watanabe, S., Mancuso, J., Green, D.R., and Sánchez Alvarado, A. (2010). Cell death and tissue remodeling in planarian regeneration. Dev Biol *338*, 76-85.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Plusquin, M., DeGheselle, O., Cuypers, A., Geerdens, E., Van Roten, A., Artois, T., and Smeets, K. (2012). Reference genes for qPCR assays in toxic metal and salinity stress in two flatworm model organisms. Ecotoxicology *21*, 475-484.

Poynton, H.C., and Vulpe, C.D. (2009). Ecotoxicogenomics: Emerging Technologies for Emerging Contaminants1. JAWRA Journal of the American Water Resources Association *45*, 83-96.

Rawls, S.M., Karaca, F., Madhani, I., Bhojani, V., Martinez, R.L., Abou-Gharbia, M., and Raffa, R.B. (2010). β-lactamase inhibitors display anti-seizure properties in an invertebrate assay. Neuroscience *169*, 1800-04.

Reitzel, A.M., Sullivan, J.C., Traylor-Knowles, N., and Finnerty, J.R. (2008). Genomic survey of candidate stress-response genes in the estuarine anemone Nematostella vectensis. Biol Bull *214*, 233-254.

Rieger, R.M., Gehlen, M., Haszprunar, G., Holmlund, M., Legniti, A., Salvenmoser, W., and Tyler, S. (1988). Laboratory cultures of marine Macrostomida (Turbellaria). Fortschr Zool *36*, 523.

Rink, J.C. (2013). Stem cell systems and regeneration in planaria. Dev Genes Evol *223*, 67-84.

Romero, R., and Baguñà, J. (1991). Quantitative cellular analysis of growth and reproduction in freshwater planarians (Turbellaria; Tricladida). I. A cellular description of the intact organism. Invertebrate reproduction & development *19*, 157-165.

Schaeffer, D.J. (1993). Planarians as a model system for in vivo teratogenesis studies. Qual Assur *2*, 265-318.

Shock, B.C., Foran, C.M., and Stueckle, T.A. (2009). Effects of salinity stress on survival, metabolism, limb regeneration, and ecdysis in Uca pugnax. Journal of Crustacean Biology *29*, 293-301.

Simanov, D., Mellaart-Straver, I., Sormacheva, I., and Berezikov, E. (2012). The Flatworm Macrostomum lignano Is a Powerful Model Organism for Ion Channel and Stem Cell Research. Stem Cells Int *2012*, 167265.

Simpson, J.T., and Durbin, R. (2012). Efficient de novo assembly of large genomes using compressed data structures. Genome Res *22*, 549-556.

Snape, J.R., Maund, S.J., Pickford, D.B., and Hutchinson, T.H. (2004). Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. Aquat Toxicol *67*, 143-154.

Steinberg, C.E., Stürzenbaum, S.R., and Menzel, R. (2008). Genes and environment - striking the fine balance between sophisticated biomonitoring and true functional environmental genomics. Sci Total Environ *400*, 142-161.

Stueckle, T.A., Shock, B., and Foran, C.M. (2009). Multiple stressor effects of methoprene, permethrin, and salinity on limb regeneration and molting in the mud fiddler crab (Uca pugnax). Environ Toxicol Chem *28*, 2348-359.

Tuteja, N. (2007). Mechanisms of high salinity tolerance in plants. Methods Enzymol *428*, 419-438.

Wang, H., Zhang, M., Guo, R., Shi, D., Liu, B., Lin, X., and Yang, C. (2012). Effects of salt stress on ion balance and nitrogen metabolism of old and young leaves in rice (Oryza sativa L.). BMC plant biology *12*, 194.

Wang, L., Showalter, A., and Ungar, I. (1997). Effect of salinity on growth, ion content, and cell wall chemistry in Atriplex prostrata (Chenopodiaceae). Am J Bot *84*, 1247.

Zhang, H., Han, B., Wang, T., Chen, S., Li, H., Zhang, Y., and Dai, S. (2012). Mechanisms of plant salt response: insights from proteomics. J Proteome Res *11*, 49-67.

SUPPLEMENTARY MATERIAL

Supplementary Table S1. Enrichment of GO Terms in genes up-regulated more than two-fold at 1h in low salinity compared to control.

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0006811	ion transport	3.30E-08	3.39E-04	1.87	AQP9; CAV3; CALCRL
GO:0016053	organic acid biosynthetic process	7.45E-07	3.83E-03	2.64	
GO:0046394	carboxylic acid biosynthetic process	7.45E-07	2.55E-03	2.64	
GO:0044283	small molecule biosynthetic process	8.00E-07	2.06E-03	2.32	
GO:0008652	cellular amino acid biosynthetic process	8.88E-07	1.82E-03	3.54	
GO:0006820	anion transport	1.19E-06	2.04E-03	2.38	TRPC4; VDAC2; CLIC1
GO:0015849	organic acid transport	1.68E-06	2.47E-03	2.76	
GO:0046942	carboxylic acid transport	1.68E-06	2.16E-03	2.76	
GO:0015074	DNA integration	2.78E-06	3.18E-03	10.76	
GO:0055085	transmembrane transport	7.04E-06	7.24E-03	1.79	ASIC3; ANO7; AQP9
GO:0019752	carboxylic acid metabolic process	1.44E-05	1.35E-02	1.72	
GO:0015711	organic anion transport	1.52E-05	1.30E-02	2.29	SYT1; STX1A
GO:1901607	alpha-amino acid biosynthetic process	2.60E-05	2.05E-02	3.73	
GO:0071320	cellular response to cAMP	3.17E-05	2.33E-02	5.74	
GO:0006082	organic acid metabolic process	3.31E-05	2.27E-02	1.64	
GO:0009612	response to mechanical stimulus	3.33E-05	2.14E-02	3.01	
GO:0051049	regulation of transport	3.42E-05	2.06E-02	1.61	
GO:0006865	amino acid transport	4.64E-05	2.65E-02	2.94	
GO:0044765	single-organism transport	5.47E-05	2.96E-02	1.35	
GO:0043436	oxoacid metabolic process	6.90E-05	3.55E-02	1.62	
GO:0051591	response to cAMP	7.95E-05	3.89E-02	3.85	
GO:0060341	regulation of cellular localization	8.89E-05	4.15E-02	1.71	
GO:0014070	response to organic cyclic compound	9.38E-05	4.19E-02	1.88	
GO:0051384	response to glucocorticoid	1.23E-04	5.28E-02	3.44	
GO:0031960	response to corticosteroid	1.84E-04	7.58E-02	3.31	
GO:0006810	Transport	1.85E-04	7.32E-02	1.27	
GO:0071705	nitrogen compound transport	1.85E-04	7.05E-02	1.87	
GO:0044699	single-organism process	1.93E-04	7.10E-02	1.09	
GO:0050982	detection of mechanical stimulus	1.99E-04	7.05E-02	4.59	
GO:0051234	establishment of localization	2.67E-04	9.16E-02	1.26	
GO:0048545	response to steroid hormone	3.04E-04	1.01E-01	2.33	
GO:0046683	response to organophosphorus	3.23E-04	1.04E-01	3.13	
GO:0006641	triglyceride metabolic process	3.86E-04	1.20E-01	3.07	
GO:0014074	response to purine-containing compound	4.56E-04	1.38E-01	2.87	
GO:0021510	spinal cord development	5.26E-04	1.54E-01	6.52	
GO:0006812	cation transport	5.37E-04	1.53E-01	1.7	NALCN; CHRNA4
GO:0032787	monocarboxylic acid metabolic process	5.62E-04	1.56E-01	1.86	
GO:0009628	response to abiotic stimulus	5.81E-04	1.57E-01	1.56	
GO:0006638	neutral lipid metabolic process	6.38E-04	1.68E-01	2.92	
GO:0006639	acylglycerol metabolic process	6.38E-04	1.64E-01	2.92	
GO:0090257	regulation of muscle system process	6.71E-04	1.68E-01	2.64	
GO:1901605	alpha-amino acid metabolic process	8.17E-04	2.00E-01	2.08	
GO:0043649	dicarboxylic acid catabolic process	8.49E-04	2.03E-01	5.98	
GO:0009070	serine family amino acid biosynthetic process	8.49E-04	1.98E-01	5.98	
GO:0043648	dicarboxylic acid metabolic process	9.55E-04	2.18E-01	2.66	
Molecular Function GO Terms					
GO:0015291	secondary active transmembrane transporter activity	4.90E-12	1.64E-08	3.72	
GO:0008509	anion transmembrane transporter activity	1.97E-11	3.31E-08	3.34	
GO:0015293	symporter activity	8.40E-11	9.39E-08	4.17	
GO:0022804	active transmembrane transporter activity	5.01E-09	4.20E-06	2.71	
GO:0022857	transmembrane transporter activity	5.81E-09	3.90E-06	1.92	
GO:0015075	ion transmembrane transporter activity	1.65E-08	9.22E-06	1.97	
GO:0015296	anion:cation symporter activity	2.10E-08	1.01E-05	5.91	

Supplementary Table S1. Enrichment of GO Terms in genes up-regulated more than two-fold at 1h in low salinity compared to control. (*Continued*)

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
GO:0046943	carboxylic acid transmembrane transporter activity	2.82E-08	1.18E-05	3.96	
GO:0005342	organic acid transmembrane transporter activity	4.64E-08	1.73E-05	3.86	
GO:0008514	organic anion transmembrane transporter activity	5.82E-08	1.95E-05	3.44	
GO:0005215	transporter activity	7.05E-08	2.15E-05	1.73	
GO:0022891	substrate-specific transmembrane transporter activity	1.10E-07	3.07E-05	1.87	
GO:0022892	substrate-specific transporter activity	1.29E-07	3.32E-05	1.8	
GO:0005283	sodium:amino acid symporter activity	7.39E-07	1.77E-04	12.3	
GO:0015294	solute:cation symporter activity	1.01E-06	2.26E-04	4.03	
GO:0005416	cation:amino acid symporter activity	7.85E-06	1.65E-03	9.56	
GO:0005343	organic acid:sodium symporter activity	1.86E-05	3.66E-03	5.38	
GO:0015108	chloride transmembrane transporter activity	2.05E-05	3.82E-03	3.81	
GO:0015297	antiporter activity	3.16E-05	5.57E-03	3.91	
GO:0005509	calcium ion binding	3.47E-05	5.81E-03	1.78	
GO:0015081	sodium ion transmembrane transporter activity	3.91E-05	6.24E-03	2.87	
GO:0015103	inorganic anion transmembrane transporter activity	4.35E-05	6.62E-03	3.21	
GO:0008324	cation transmembrane transporter activity	8.31E-05	1.21E-02	1.74	
GO:0015377	cation:chloride symporter activity	1.11E-04	1.54E-02	11.48	
GO:0015370	solute:sodium symporter activity	2.52E-04	3.38E-02	3.68	
GO:0015171	amino acid transmembrane transporter activity	3.63E-04	4.69E-02	3.29	
GO:0046873	metal ion transmembrane transporter activity	3.79E-04	4.71E-02	1.84	
GO:0005154	epidermal growth factor receptor binding	5.26E-04	6.29E-02	6.52	
GO:0008271	secondary active sulfate transmembrane transporter activity	6.91E-04	7.99E-02	8.2	
GO:1901682	sulfur compound transmembrane transporter activity	7.13E-04	7.97E-02	5.06	
GO:0022803	passive transmembrane transporter activity	8.20E-04	8.87E-02	1.8	
GO:0015267	channel activity	8.20E-04	8.59E-02	1.8	
GO:0008028	monocarboxylic acid transmembrane transporter activity	9.39E-04	9.55E-02	4.18	
GO:0016791	phosphatase activity	9.72E-04	9.59E-02	1.98	

Supplementary Table S2. Enrichment of GO Terms in genes down-regulated more than two-fold at 1h in low salinity compared to control.

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0007165	signal transduction	1.20E-08	1.23E-04	1.52	BRCA1; NOTCH1; TNK2
GO:0097190	apoptotic signaling pathway	2.25E-07	1.15E-03	3.19	CASP3; CASP10; CASP9
GO:0001508	regulation of action potential	9.32E-07	3.19E-03	6.07	
GO:0070848	response to growth factor	1.88E-06	4.83E-03	2.29	
GO:0051716	cellular response to stimulus	1.97E-06	4.04E-03	1.36	
GO:0050896	response to stimulus	2.46E-06	4.22E-03	1.3	
GO:0048011	neurotrophin TRK receptor signaling pathway	2.92E-06	4.29E-03	2.79	
GO:0071363	cellular response to growth factor stimulus	3.22E-06	4.13E-03	2.27	
GO:0038179	neurotrophin signaling pathway	3.57E-06	4.08E-03	2.76	
GO:0007166	cell surface receptor signaling pathway	3.69E-06	3.80E-03	1.6	CDIP1; TRAF2; ITPR1
GO:0044707	single-multicellular organism process	9.07E-06	8.47E-03	1.47	
GO:0032501	multicellular organismal process	1.74E-05	1.49E-02	1.45	
GO:0045665	negative regulation of neuron differentiation	2.64E-05	2.08E-02	5.4	
GO:0044699	single-organism process	3.57E-05	2.62E-02	1.12	
GO:0043122	regulation of I-kappaB kinase/NF-kappaB signaling	5.98E-05	4.10E-02	3.05	
GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	7.14E-05	4.59E-02	3.49	
GO:0071310	cellular response to organic substance	9.32E-05	5.63E-02	1.62	
GO:0023051	regulation of signaling	1.12E-04	6.42E-02	1.48	
GO:0070887	cellular response to chemical stimulus	1.14E-04	6.17E-02	1.56	
GO:0010646	regulation of cell communication	1.23E-04	6.32E-02	1.47	
GO:0051865	protein autoubiquitination	1.44E-04	7.05E-02	5.71	
GO:0007154	cell communication	1.63E-04	7.63E-02	1.8	
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	1.89E-04	8.46E-02	1.99	NGFR; INSRR
GO:2001236	regulation of extrinsic apoptotic signaling pathway	1.91E-04	8.19E-02	4.27	TRAF1; SCRT2
GO:0060341	regulation of cellular localization	1.92E-04	7.91E-02	1.82	
GO:0051960	regulation of nervous system development	1.97E-04	7.78E-02	2.04	
GO:0050767	regulation of neurogenesis	2.12E-04	8.08E-02	2.1	
GO:0090002	establishment of protein localization to plasma membrane	2.28E-04	8.37E-02	4.66	
GO:0006811	ion transport	2.48E-04	8.80E-02	1.67	CNGB1; TRPM2
GO:0086001	regulation of cardiac muscle cell action potential	2.56E-04	8.78E-02	7.85	
GO:0042391	regulation of membrane potential	2.57E-04	8.53E-02	2.61	SCN5A; KCND2
GO:0017157	regulation of exocytosis	2.90E-04	9.31E-02	3.71	
GO:0051049	regulation of transport	3.25E-04	1.01E-01	1.64	
GO:0002831	regulation of response to biotic stimulus	3.28E-04	9.91E-02	3.99	
GO:0030216	keratinocyte differentiation	3.43E-04	1.01E-01	4.41	
GO:0009966	regulation of signal transduction	4.30E-04	1.23E-01	1.47	MAP2K1; SOX2
GO:0044700	single organism signaling	4.50E-04	1.25E-01	1.83	
GO:0023052	signaling	4.50E-04	1.22E-01	1.83	
GO:0042221	response to chemical	4.69E-04	1.24E-01	1.39	
GO:0006816	calcium ion transport	5.22E-04	1.34E-01	2.75	
GO:0044767	single-organism developmental process	5.35E-04	1.34E-01	1.3	
GO:0086036	regulation of cardiac muscle cell membrane potential	5.51E-04	1.35E-01	6.8	
GO:0060627	regulation of vesicle-mediated transport	5.60E-04	1.34E-01	2.44	
GO:0008631	intrinsic apoptotic signaling pathway in response to oxidative stress	5.89E-04	1.38E-01	9.07	
GO:0044763	single-organism cellular process	6.24E-04	1.43E-01	1.12	
GO:0032879	regulation of localization	9.07E-04	2.03E-01	1.49	
GO:0070838	divalent metal ion transport	9.22E-04	2.02E-01	2.6	
GO:0045664	regulation of neuron differentiation	9.51E-04	2.04E-01	2.07	
Molecular Function GO Terms					
GO:0031996	thioesterase binding	1.80E-07	6.05E-04	12.99	
GO:0022803	passive transmembrane transporter activity	2.45E-06	4.12E-03	2.48	
GO:0015267	channel activity	2.45E-06	2.74E-03	2.48	
GO:0005509	calcium ion binding	4.26E-06	3.57E-03	2.09	
GO:0022838	substrate-specific channel activity	4.67E-06	3.13E-03	2.45	

Supplementary Table S2. Enrichment of GO Terms in genes down-regulated more than two-fold at 1h in low salinity compared to control. (*Continued*)

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
GO:0005216	ion channel activity	8.82E-06	4.93E-03	2.42	
GO:0050998	nitric-oxide synthase binding	1.37E-05	6.56E-03	12.75	
GO:0005261	cation channel activity	4.47E-05	1.87E-02	2.5	
GO:0015075	ion transmembrane transporter activity	5.46E-05	2.03E-02	1.82	
GO:0022857	transmembrane transporter activity	1.20E-04	4.04E-02	1.71	
GO:0005215	transporter activity	1.41E-04	4.31E-02	1.61	
GO:0016881	acid-amino acid ligase activity	1.77E-04	4.96E-02	2.24	
GO:0022834	ligand-gated channel activity	2.01E-04	5.18E-02	3.01	
GO:0015276	ligand-gated ion channel activity	2.01E-04	4.81E-02	3.01	
GO:0004842	ubiquitin-protein ligase activity	2.01E-04	4.49E-02	2.31	
GO:0046873	metal ion transmembrane transporter activity	2.40E-04	5.04E-02	2.08	
GO:0022891	substrate-specific transmembrane transporter activity	2.55E-04	5.04E-02	1.7	
GO:0043169	cation binding	3.50E-04	6.52E-02	1.27	
GO:0050699	WW domain binding	3.83E-04	6.76E-02	7.29	
GO:0015081	sodium ion transmembrane transporter activity	4.11E-04	6.90E-02	2.95	
GO:0005231	excitatory extracellular ligand-gated ion channel activity	4.17E-04	6.65E-02	4.3	
GO:0019787	small conjugating protein ligase activity	4.39E-04	6.69E-02	2.19	
GO:0004889	acetylcholine-activated cation-selective channel activity	5.51E-04	8.04E-02	6.8	
GO:0050997	quaternary ammonium group binding	5.51E-04	7.71E-02	6.8	
GO:0042165	neurotransmitter binding	5.51E-04	7.40E-02	6.8	
GO:0044389	small conjugating protein ligase binding	5.76E-04	7.43E-02	2.72	
GO:0031625	ubiquitin protein ligase binding	5.76E-04	7.15E-02	2.72	
GO:0008324	cation transmembrane transporter activity	6.76E-04	8.09E-02	1.77	
GO:0043167	ion binding	7.37E-04	8.52E-02	1.18	
GO:0022892	substrate-specific transporter activity	7.59E-04	8.49E-02	1.59	
GO:0016879	ligase activity, forming carbon-nitrogen bonds	9.24E-04	1.00E-01	2	
GO:0042166	acetylcholine binding	9.44E-04	9.89E-02	8.16	

Supplementary Table S3. Enrichment of GO Terms in genes up-regulated more than two-fold at 60 hours of starvation compared to control (first hour of starvation).

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0045022	early endosome to late endosome transport	3.12E-04	1.00E+00	21.34	FLNA; RAB7A; HOOK3
GO:0042552	myelination	3.55E-04	1.00E+00	11.59	
GO:0008366	axon ensheathment	4.10E-04	1.00E+00	11.18	
GO:0007272	ensheathment of neurons	4.10E-04	1.00E+00	11.18	
GO:0019433	triglyceride catabolic process	4.12E-04	8.46E-01	19.56	PRKACA
GO:0006006	glucose metabolic process	4.24E-04	7.26E-01	5.12	PKM, ALDOA
GO:0034088	maintenance of mitotic sister chromatid cohesion	4.82E-04	7.07E-01	52.17	
GO:0034086	maintenance of sister chromatid cohesion	4.82E-04	6.19E-01	52.17	
GO:0046461	neutral lipid catabolic process	5.30E-04	6.05E-01	18.06	
GO:0046464	acylglycerol catabolic process	5.30E-04	5.45E-01	18.06	
GO:0006641	triglyceride metabolic process	7.12E-04	6.65E-01	6.99	ELOVL4
GO:0006638	neutral lipid metabolic process	9.05E-04	7.75E-01	6.63	
GO:0006639	acylglycerol metabolic process	9.05E-04	7.15E-01	6.63	
Molecular Function GO Terms					
GO:0017137	Rab GTPase binding	1.63E-04	5.46E-01	9.54	
GO:0017016	Ras GTPase binding	4.75E-04	7.96E-01	5.03	
GO:0031267	small GTPase binding	7.26E-04	8.12E-01	4.68	
GO:0015379	potassium:chloride symporter activity	9.56E-04	8.01E-01	39.13	
GO:0030306	ADP-ribosylation factor binding	9.56E-04	6.41E-01	39.13	
GO:0036033	mediator complex binding	9.56E-04	5.34E-01	39.13	

Supplementary Table S4. Enrichment of GO Terms in genes down-regulated more than two-fold at 60 hours of starvation compared to control (first hour of starvation).

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0035725	sodium ion transmembrane transport	1.71E-06	1.75E-02	21.41	
GO:0006811	ion transport	6.26E-05	3.21E-01	2.2	
GO:0048747	muscle fiber development	8.05E-05	2.76E-01	10.71	
GO:0006814	Sodium ion transport	9.18E-05	2.36E-01	6.47	
GO:0015672	monovalent inorganic cation transport	1.10E-04	2.26E-01	3.65	
GO:0055001	muscle cell development	2.69E-04	4.61E-01	6.57	
GO:0003008	system process	2.92E-04	4.28E-01	2.08	PNLIP; GRM5
GO:0006812	cation transport	3.98E-04	5.11E-01	2.45	
GO:0071436	Sodium ion export	4.48E-04	5.12E-01	47.11	
GO:0045008	depyrimidination	4.48E-04	4.61E-01	47.11	
GO:0050968	detection of chemical stimulus involved in sensory perception of pain	4.48E-04	4.19E-01	47.11	
GO:0009223	pyrimidine deoxyribonucleotide catabolic process	4.48E-04	3.84E-01	47.11	TDG, NTHL1
GO:0019932	Second-messenger-mediated signaling	4.72E-04	3.73E-01	5	DMD, ITPR1
GO:0030001	metal ion transport	5.12E-04	3.76E-01	2.67	
GO:0016042	lipid catabolic process	6.56E-04	4.49E-01	3.44	
GO:0097150	neuronal stem cell maintenance	7.20E-04	4.63E-01	15.7	NOTCH1
GO:0071320	cellular response to cAMP	7.31E-04	4.42E-01	9.42	
GO:0055006	cardiac cell development	7.31E-04	4.17E-01	9.42	
Molecular Function GO Terms					
GO:0005509	calcium ion binding	3.93E-08	1.32E-04	3.21	
GO:0015081	sodium ion transmembrane transporter activity	1.17E-04	1.96E-01	4.71	
GO:0015297	antiporter activity	3.06E-04	3.42E-01	6.42	
GO:0015291	secondary active transmembrane transporter activity	5.84E-04	4.90E-01	3.49	
GO:0004622	lysophospholipase activity	7.20E-04	4.83E-01	15.7	

Supplementary Table S5. Enrichment of GO Terms in genes down-regulated more than two-fold at 60 hours in low salinity compared to 60 hours of starvation.

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0006811	ion transport	7.17E-10	7.37E-06	2.34	NALCN; NOX1; CLCA1
GO:0055085	transmembrane transport	5.54E-09	2.84E-05	2.45	KCNIP2; ASIC5;
GO:0006820	anion transport	1.05E-08	3.61E-05	3.32	CLIC1; CTNS; CLCA1
GO:0015711	organic anion transport	2.16E-07	5.54E-04	3.21	CA7; HTT
GO:0072348	sulfur compound transport	4.07E-07	8.36E-04	10.17	
GO:0043436	oxoacid metabolic process	1.44E-06	2.47E-03	2.04	ADC; HAL; SMS; CDO1
GO:0019752	carboxylic acid metabolic process	1.50E-06	2.20E-03	2.1	GLUL; GLUD1; PSAT1
GO:0006082	organic acid metabolic process	2.26E-06	2.90E-03	2.01	AGXT2; HAGH; MTAP
GO:0044765	single-organism transport	3.28E-06	3.74E-03	1.54	
GO:1901605	alpha-amino acid metabolic process	4.24E-06	4.35E-03	3.18	PDPR; DDAH2; PYCR2
GO:0015849	organic acid transport	1.06E-05	9.94E-03	3.22	
GO:0046942	carboxylic acid transport	1.06E-05	9.11E-03	3.22	
GO:0015698	inorganic anion transport	1.20E-05	9.51E-03	6.9	
GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	1.22E-05	8.97E-03	4.13	
GO:0016053	organic acid biosynthetic process	1.24E-05	8.46E-03	2.96	
GO:0046394	carboxylic acid biosynthetic process	1.24E-05	7.93E-03	2.96	
GO:0006821	chloride transport	1.76E-05	1.07E-02	9.66	
GO:0006520	cellular amino acid metabolic process	1.95E-05	1.11E-02	2.36	
GO:0035725	sodium ion transmembrane transport	4.45E-05	2.41E-02	10.98	
GO:0006090	pyruvate metabolic process	4.97E-05	2.55E-02	6.76	
GO:0008652	cellular amino acid biosynthetic process	6.87E-05	3.36E-02	3.76	
GO:0006596	polyamine biosynthetic process	9.17E-05	4.28E-02	13.8	
GO:0006595	polyamine metabolic process	1.16E-04	5.17E-02	9.29	
GO:0043122	regulation of I-kappaB kinase/NF-kappaB signaling	1.25E-04	5.35E-02	3.16	
GO:1901607	alpha-amino acid biosynthetic process	1.81E-04	7.42E-02	4.35	
GO:0006814	sodium ion transport	2.11E-04	8.35E-02	4.26	
GO:1900744	regulation of p38MAPK cascade	2.73E-04	1.04E-01	18.12	
GO:0044283	small molecule biosynthetic process	2.89E-04	1.06E-01	2.26	
GO:1901606	alpha-amino acid catabolic process	3.81E-04	1.35E-01	3.36	
GO:0006810	transport	3.92E-04	1.34E-01	1.34	
GO:0050982	detection of mechanical stimulus	4.39E-04	1.45E-01	5.8	
GO:0008272	sulfate transport	4.98E-04	1.60E-01	9.66	
GO:0042401	cellular biogenic amine biosynthetic process	4.98E-04	1.55E-01	9.66	
GO:0051234	establishment of localization	5.59E-04	1.69E-01	1.33	
GO:0008016	regulation of heart contraction	5.66E-04	1.66E-01	3.45	
GO:0006865	amino acid transport	5.86E-04	1.67E-01	3.2	
GO:0055117	regulation of cardiac muscle contraction	6.84E-04	1.90E-01	5.37	
GO:0044281	small molecule metabolic process	7.21E-04	1.95E-01	1.39	
GO:0007271	synaptic transmission, cholinergic	7.57E-04	1.99E-01	8.78	
GO:0009309	amine biosynthetic process	7.57E-04	1.94E-01	8.78	
GO:0009063	cellular amino acid catabolic process	8.33E-04	2.09E-01	2.9	
GO:0034220	ion transmembrane transport	8.37E-04	2.05E-01	2.24	
GO:0009069	serine family amino acid metabolic process	8.40E-04	2.01E-01	5.18	
GO:0009084	glutamine family amino acid biosynthetic process	8.52E-04	1.99E-01	6.36	
GO:0022414	reproductive process	9.67E-04	2.21E-01	1.77	CDKL2; HSPA2; SPA17
Molecular Function GO Terms					
GO:0008509	anion transmembrane transporter activity	6.08E-13	2.04E-09	4.59	
GO:0015291	secondary active transmembrane transporter activity	8.98E-13	1.51E-09	5.01	
GO:0022857	transmembrane transporter activity	2.81E-12	3.14E-09	2.55	
GO:0015075	ion transmembrane transporter activity	1.94E-11	1.63E-08	2.62	
GO:0022891	substrate-specific transmembrane transporter activity	8.51E-11	5.71E-08	2.49	
GO:0008514	organic anion transmembrane transporter activity	5.35E-10	2.99E-07	5.07	
GO:0015293	symporter activity	8.90E-10	4.26E-07	5.19	
GO:0022804	active transmembrane transporter activity	2.89E-09	1.21E-06	3.42	
GO:0005215	transporter activity	4.42E-09	1.65E-06	2.09	
GO:0022892	substrate-specific transporter activity	4.58E-09	1.54E-06	2.23	
GO:0046943	carboxylic acid transmembrane transporter activity	6.09E-08	1.86E-05	5.09	

Supplementary Table S5. Enrichment of GO Terms in genes down-regulated more than two-fold at 60 hours in low salinity compared to 60 hours of starvation. (*Continued*)

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
GO:0005342	organic acid transmembrane transporter activity	8.96E-08	2.50E-05	4.95	
GO:0015081	sodium ion transmembrane transporter activity	1.27E-07	3.27E-05	4.56	
GO:1901682	sulfur compound transmembrane transporter activity	2.68E-06	6.42E-04	9.95	
GO:0015294	solute:cation symporter activity	2.79E-06	6.25E-04	5.09	
GO:0015103	inorganic anion transmembrane transporter activity	2.85E-06	5.97E-04	4.69	
GO:0015108	chloride transmembrane transporter activity	3.75E-06	7.39E-04	5.42	
GO:0008324	cation transmembrane transporter activity	4.02E-06	7.48E-04	2.22	
GO:0016645	oxidoreductase activity, acting on the CH-NH group of donors	9.59E-06	1.69E-03	8.45	
GO:0046873	metal ion transmembrane transporter activity	1.34E-05	2.25E-03	2.46	
GO:0015301	anion:anion antiporter activity	1.76E-05	2.82E-03	9.66	
GO:0015370	solute:sodium symporter activity	2.29E-05	3.49E-03	5.57	
GO:0015297	antiporter activity	6.36E-05	9.28E-03	4.94	
GO:0046914	transition metal ion binding	6.79E-05	9.49E-03	1.65	
GO:0005216	ion channel activity	7.62E-05	1.02E-02	2.37	
GO:0008271	secondary active sulfate transmembrane transporter activity	9.17E-05	1.18E-02	13.8	
GO:0022838	substrate-specific channel activity	1.05E-04	1.30E-02	2.32	
GO:0022803	passive transmembrane transporter activity	1.42E-04	1.71E-02	2.27	
GO:0015267	channel activity	1.42E-04	1.65E-02	2.27	
GO:0005328	neurotransmitter:sodium symporter activity	1.74E-04	1.95E-02	8.63	
GO:0005452	inorganic anion exchanger activity	1.77E-04	1.92E-02	12.08	
GO:0005254	chloride channel activity	1.93E-04	2.02E-02	4.83	
GO:0004499	N,N-dimethylaniline monooxygenase activity	2.73E-04	2.78E-02	18.12	
GO:0005253	anion channel activity	2.75E-04	2.71E-02	4.6	
GO:0015116	sulfate transmembrane transporter activity	3.09E-04	2.96E-02	10.74	
GO:0022890	inorganic cation transmembrane transporter activity	3.39E-04	3.16E-02	2.04	
GO:0015296	anion:cation symporter activity	4.03E-04	3.65E-02	4.97	
GO:0016491	oxidoreductase activity	4.97E-04	4.38E-02	1.82	
GO:0030507	spectrin binding	4.98E-04	4.28E-02	9.66	
GO:0015171	amino acid transmembrane transporter activity	7.07E-04	5.93E-02	4.03	
GO:0044325	ion channel binding	7.07E-04	5.79E-02	4.03	
GO:0043236	laminin binding	7.57E-04	6.05E-02	8.78	
GO:0005326	neurotransmitter transporter activity	8.52E-04	6.64E-02	6.36	
GO:0008270	zinc ion binding	9.57E-04	7.30E-02	1.58	

Supplementary Table S6. Enrichment of GO Terms in genes up-regulated more than two-fold at 60 hours in low salinity compared to 60 hours of starvation.

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0030049	muscle filament sliding	3.13E-07	3.21E-03	7.17	MYH3; MYH2; TTN
GO:0033275	actin-myosin filament sliding	3.13E-07	1.61E-03	7.17	TPM3, MYH7
GO:0006811	ion transport	1.31E-06	4.47E-03	1.79	TRPA1; ABCA1
GO:0003008	system process	1.46E-06	3.76E-03	1.8	CASP3
GO:0032501	multicellular organismal process	3.03E-06	6.22E-03	1.43	
GO:0044707	single-multicellular organism process	3.67E-06	6.29E-03	1.43	
GO:0070252	actin-mediated cell contraction	4.48E-06	6.57E-03	5.64	
GO:0006928	cellular component movement	4.74E-06	6.08E-03	1.67	
GO:0007155	cell adhesion	9.26E-06	1.06E-02	1.88	
GO:0022610	biological adhesion	1.06E-05	1.09E-02	1.87	
GO:0030048	actin filament-based movement	2.57E-05	2.40E-02	4.34	
GO:0006816	calcium ion transport	8.24E-05	7.06E-02	2.73	
GO:0008016	regulation of heart contraction	1.02E-04	8.08E-02	3.16	
GO:0070838	divalent metal ion transport	1.73E-04	1.27E-01	2.58	

Supplementary Table S6. Enrichment of GO Terms in genes up-regulated more than two-fold at 60 hours in low salinity compared to 60 hours of starvation. (*Continued*)

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
GO:0072511	divalent inorganic cation transport	1.95E-04	1.34E-01	2.56	
GO:0032414	positive regulation of ion transmembrane transporter activity	2.10E-04	1.35E-01	5.26	
GO:1902533	positive regulation of intracellular signal transduction	2.39E-04	1.44E-01	1.85	
GO:0044700	single organism signaling	2.58E-04	1.47E-01	1.75	CASP3; CASP8; NOTCH1
GO:0023052	signaling	2.58E-04	1.40E-01	1.75	
GO:0007154	cell communication	2.58E-04	1.33E-01	1.67	
GO:0002027	regulation of heart rate	3.11E-04	1.52E-01	3.95	
GO:0006873	cellular ion homeostasis	5.71E-04	2.67E-01	2.02	
GO:0007220	Notch receptor processing	6.07E-04	2.71E-01	5.26	
GO:0032411	positive regulation of transporter activity	6.97E-04	2.98E-01	4.42	
GO:0050982	detection of mechanical stimulus	6.97E-04	2.86E-01	4.42	
GO:0046903	secretion	7.07E-04	2.79E-01	1.84	
GO:0030003	cellular cation homeostasis	7.20E-04	2.74E-01	2.03	
GO:0044765	single-organism transport	8.28E-04	3.04E-01	1.3	
GO:0043269	regulation of ion transport	8.33E-04	2.95E-01	2.08	
Molecular Function GO Terms					
GO:0005509	calcium ion binding	2.89E-11	9.69E-08	2.42	
GO:0022857	transmembrane transporter activity	2.51E-08	4.20E-05	1.93	
GO:0022891	substrate-specific transmembrane transporter activity	5.93E-08	6.64E-05	1.94	
GO:0015075	ion transmembrane transporter activity	1.26E-07	1.06E-04	1.95	
GO:0005215	transporter activity	8.97E-07	6.01E-04	1.7	
GO:0022892	substrate-specific transporter activity	1.87E-06	1.04E-03	1.76	
GO:0015085	calcium ion transmembrane transporter activity	2.23E-06	1.07E-03	3.37	
GO:0072509	divalent inorganic cation transmembrane transporter activity	2.78E-06	1.17E-03	3.1	
GO:0015291	secondary active transmembrane transporter activity	3.47E-06	1.29E-03	2.81	
GO:0022804	active transmembrane transporter activity	9.88E-06	3.31E-03	2.31	
GO:0005262	calcium channel activity	1.18E-05	3.60E-03	3.41	
GO:0008237	metallopeptidase activity	1.38E-05	3.87E-03	2.9	
GO:0008324	cation transmembrane transporter activity	1.48E-05	3.81E-03	1.88	
GO:0022803	passive transmembrane transporter activity	6.61E-05	1.58E-02	2.04	
GO:0015267	channel activity	6.61E-05	1.48E-02	2.04	
GO:0005216	ion channel activity	7.11E-05	1.49E-02	2.06	
GO:0005198	structural molecule activity	1.03E-04	2.02E-02	1.79	
GO:0022838	substrate-specific channel activity	1.05E-04	1.96E-02	2.02	
GO:0046873	metal ion transmembrane transporter activity	1.40E-04	2.47E-02	1.97	
GO:0022890	inorganic cation transmembrane transporter activity	2.60E-04	4.36E-02	1.82	
GO:0043169	cation binding	2.95E-04	4.70E-02	1.24	
GO:0008307	structural constituent of muscle	3.24E-04	4.94E-02	4.35	
GO:0005261	cation channel activity	3.27E-04	4.77E-02	2.1	
GO:0000146	microfilament motor activity	4.27E-04	5.97E-02	5.57	
GO:0046872	metal ion binding	4.63E-04	6.21E-02	1.24	
GO:0008271	secondary active sulfate transmembrane transporter activity	4.79E-04	6.18E-02	9.02	
GO:0015081	sodium ion transmembrane transporter activity	4.83E-04	6.00E-02	2.63	
GO:0008028	monocarboxylic acid transmembrane transporter activity	5.30E-04	6.35E-02	4.6	
GO:0004222	metalloendopeptidase activity	5.89E-04	6.81E-02	2.96	
GO:0043167	ion binding	8.36E-04	9.34E-02	1.16	

Supplementary Table S7. Enrichment of GO Terms in genes down-regulated more than two-fold only at 1h in low salinity (and not later) comparing to control.

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0030049	muscle filament sliding	3.13E-07	3.23E-03	7.17	TTN; MYH3; MYH7
GO:0033275	actin-myosin filament sliding	3.13E-07	1.62E-03	7.17	MYH8, TPM3
GO:0006811	ion transport	1.48E-06	5.09E-03	1.79	DRD4; TRPC6; TRPM1
GO:0003008	system process	1.76E-06	4.55E-03	1.79	
GO:0032501	multicellular organismal process	3.52E-06	7.27E-03	1.42	
GO:0044707	single-multicellular organism process	4.27E-06	7.34E-03	1.42	
GO:0070252	actin-mediated cell contraction	4.48E-06	6.61E-03	5.64	TNNI2
GO:0006928	cellular component movement	6.15E-06	7.94E-03	1.66	
GO:0007155	cell adhesion	9.93E-06	1.14E-02	1.88	FN1; DST; CHL1
GO:0022610	biological adhesion	1.13E-05	1.17E-02	1.87	
GO:0030048	actin filament-based movement	2.57E-05	2.42E-02	4.34	MYO7A
GO:0008016	regulation of heart contraction	1.02E-04	8.80E-02	3.16	
GO:0032414	positive regulation of ion transmembrane transporter activity	2.11E-04	1.67E-01	5.26	
GO:1902533	positive regulation of intracellular signal transduction	2.55E-04	1.88E-01	1.84	
GO:0006816	calcium ion transport	2.64E-04	1.82E-01	2.58	
GO:0007154	cell communication	2.73E-04	1.76E-01	1.66	
GO:0044700	single organism signaling	2.74E-04	1.66E-01	1.74	
GO:0023052	signaling	2.74E-04	1.57E-01	1.74	
GO:0002027	regulation of heart rate	3.11E-04	1.69E-01	3.94	
GO:0070838	divalent metal ion transport	5.21E-04	2.69E-01	2.44	
GO:0006873	cellular ion homeostasis	5.72E-04	2.81E-01	2.02	
GO:0072511	divalent inorganic cation transport	5.80E-04	2.72E-01	2.42	
GO:0032411	positive regulation of transporter activity	6.97E-04	3.13E-01	4.42	
GO:0050982	detection of mechanical stimulus	6.97E-04	3.00E-01	4.42	
GO:0030003	cellular cation homeostasis	7.21E-04	2.98E-01	2.03	
GO:0046903	secretion	7.54E-04	2.99E-01	1.83	
GO:0044765	single-organism transport	8.75E-04	3.35E-01	1.29	
Molecular Function GO Terms					
GO:0005509	calcium ion binding	2.90E-11	9.75E-08	2.42	
GO:0022857	transmembrane transporter activity	2.52E-08	4.23E-05	1.93	
GO:0022891	substrate-specific transmembrane transporter activity	2.82E-08	3.15E-05	1.97	
GO:0015075	ion transmembrane transporter activity	5.95E-08	5.00E-05	1.98	
GO:0015085	calcium ion transmembrane transporter activity	6.13E-07	4.12E-04	3.51	
GO:0072509	divalent inorganic cation transmembrane transporter activity	8.34E-07	4.67E-04	3.21	
GO:0005215	transporter activity	1.00E-06	4.82E-04	1.69	
GO:0022892	substrate-specific transporter activity	1.11E-06	4.65E-04	1.77	
GO:0015291	secondary active transmembrane transporter activity	3.48E-06	1.30E-03	2.81	
GO:0008324	cation transmembrane transporter activity	7.31E-06	2.46E-03	1.91	
GO:0022804	active transmembrane transporter activity	9.90E-06	3.03E-03	2.31	
GO:0005262	calcium channel activity	1.18E-05	3.32E-03	3.41	
GO:0008237	metallopeptidase activity	1.39E-05	3.58E-03	2.9	
GO:0022803	passive transmembrane transporter activity	6.62E-05	1.59E-02	2.04	
GO:0015267	channel activity	6.62E-05	1.48E-02	2.04	
GO:0046873	metal ion transmembrane transporter activity	6.62E-05	1.39E-02	2.02	
GO:0005216	ion channel activity	7.12E-05	1.41E-02	2.06	
GO:0005198	structural molecule activity	1.03E-04	1.92E-02	1.79	
GO:0022838	substrate-specific channel activity	1.05E-04	1.86E-02	2.02	
GO:0022890	inorganic cation transmembrane transporter activity	1.32E-04	2.23E-02	1.87	
GO:0015081	sodium ion transmembrane transporter activity	1.91E-04	3.06E-02	2.74	
GO:0008307	structural constituent of muscle	3.24E-04	4.95E-02	4.35	
GO:0043169	cation binding	3.25E-04	4.75E-02	1.24	
GO:0005261	cation channel activity	3.28E-04	4.59E-02	2.1	
GO:0000146	microfilament motor activity	4.28E-04	5.75E-02	5.57	
GO:0008271	secondary active sulfate transmembrane transporter activity	4.79E-04	6.19E-02	9.02	
GO:0046872	metal ion binding	4.98E-04	6.20E-02	1.23	

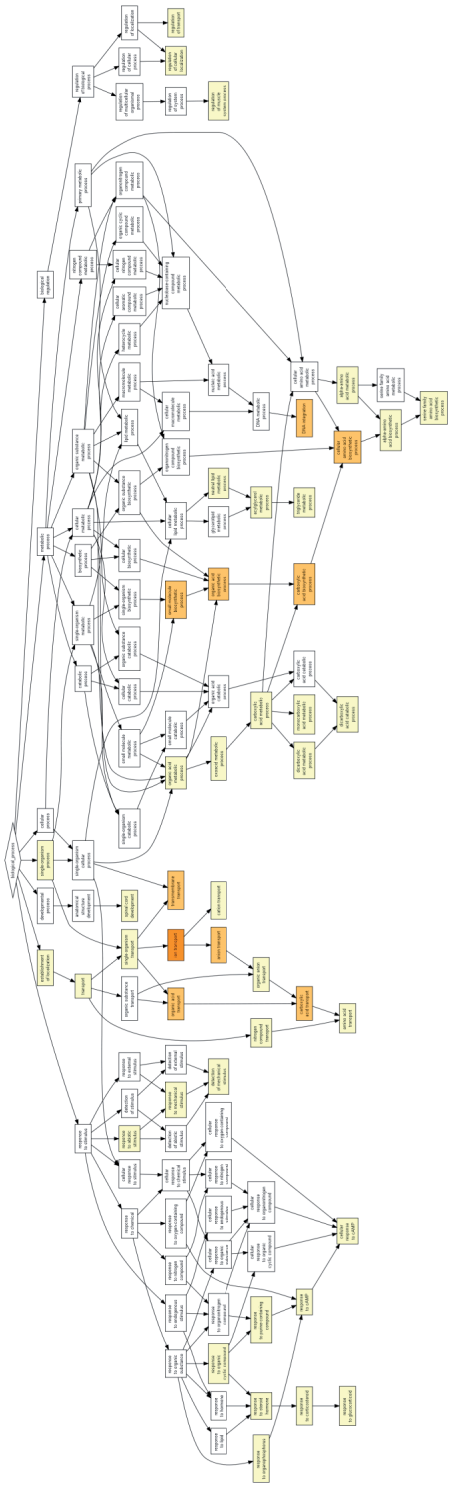
Supplementary Table S7. Enrichment of GO Terms in genes down-regulated more than two-fold only at 1h in low salinity (and not later) comparing to control. (*Continued*)

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
GO:0008028	monocarboxylic acid transmembrane transporter activity	5.31E-04	6.37E-02	4.6	
GO:0004222	metalloendopeptidase activity	5.89E-04	6.83E-02	2.96	

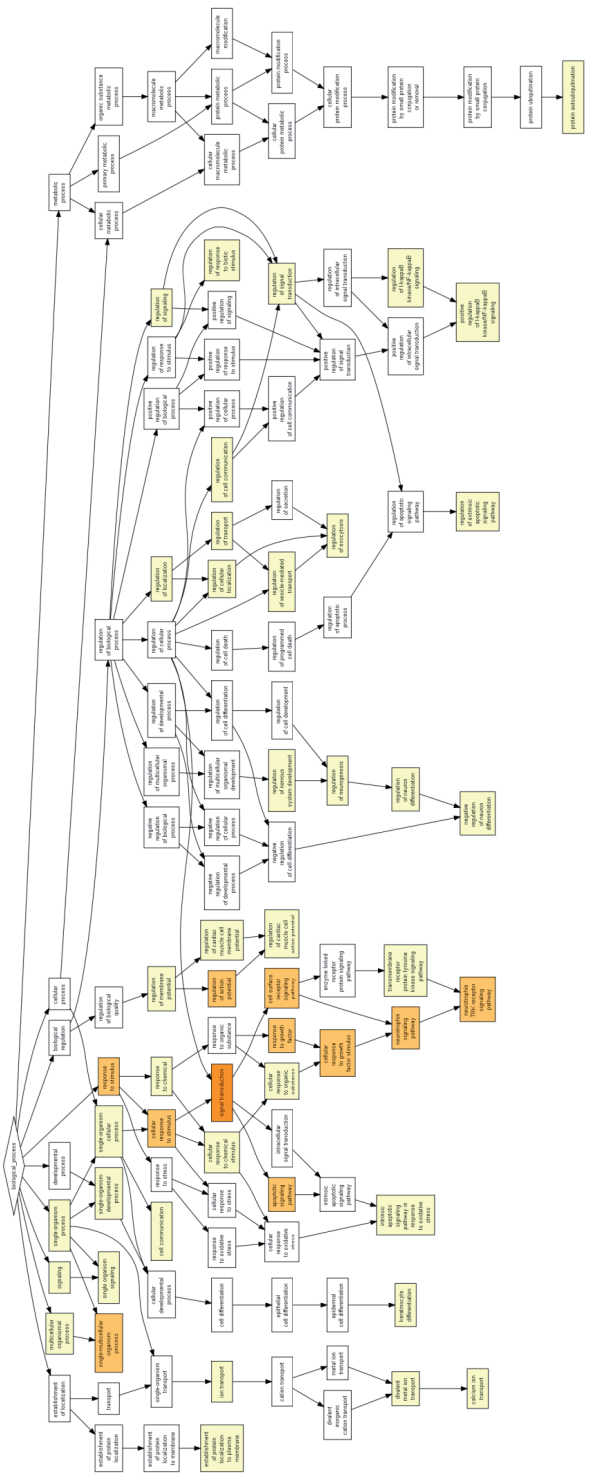
Supplementary Table S8. Enrichment of GO Terms in genes up-regulated more than two-fold only at 1h in low salinity (and not later) comparing to control.

GO Term	Description	P-value	FDR q-value	Enrichment	Gene examples
Biological Process GO Terms					
GO:0071320	cellular response to cAMP	3.55E-04	1.00E+00	11.41	EGR1; RAPGEF1
GO:0006470	protein dephosphorylation	6.55E-04	1.00E+00	4.75	PTPN1; PTPRG; PTPN11
GO:0071357	cellular response to type I interferon	7.87E-04	1.00E+00	15.56	EGR1; PTPN1
GO:0021510	spinal cord development	7.87E-04	1.00E+00	15.56	
GO:0060337	type I interferon signaling pathway	7.87E-04	1.00E+00	15.56	
GO:0032854	positive regulation of Rap GTPase activity	9.06E-04	1.00E+00	38.02	
Molecular Function GO Terms					
GO:0008131	primary amine oxidase activity	9.06E-04	1.00E+00	38.02	
GO:0017034	Rap guanyl-nucleotide exchange factor activity	9.06E-04	1.00E+00	38.02	

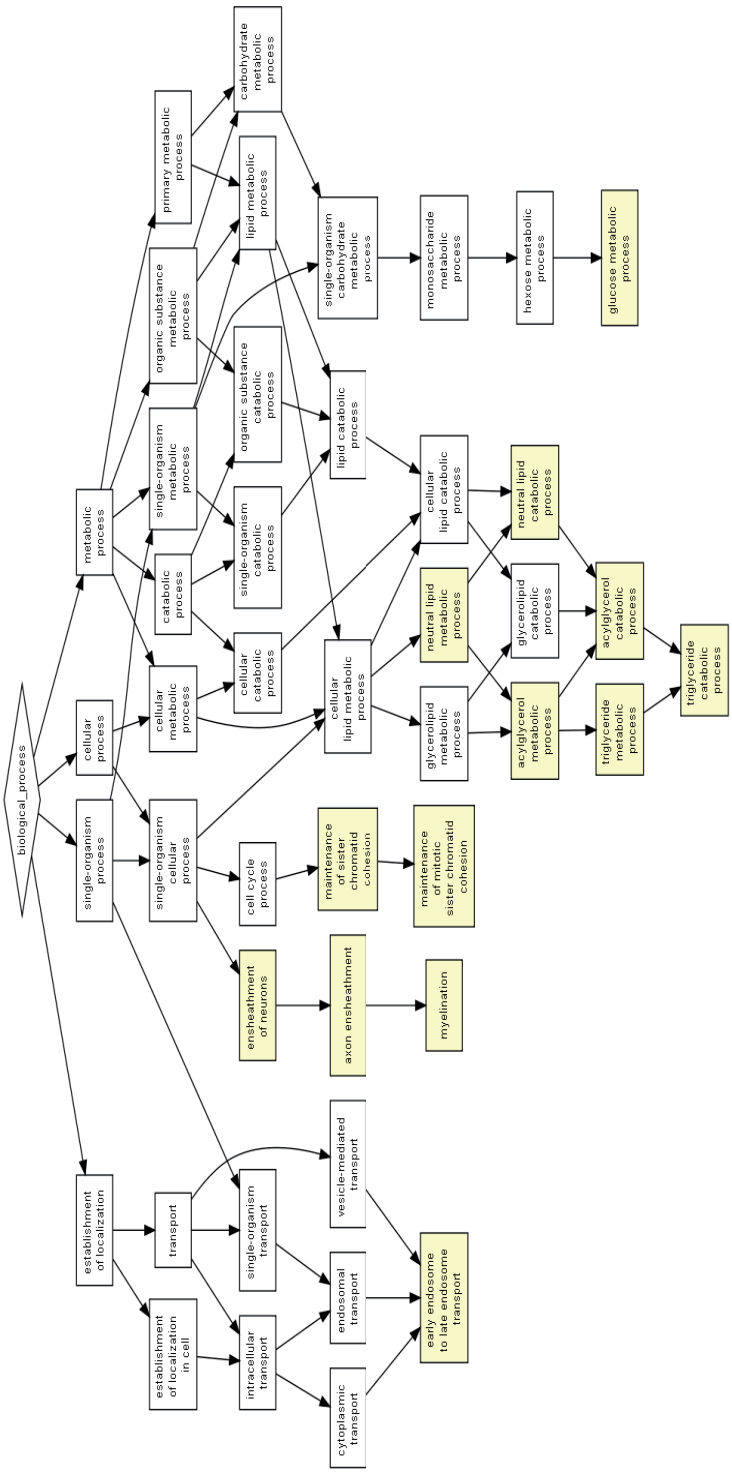
In all Supplementary Tables: "P-value" is the enrichment p-value computed according to the mHG or HG model. This p-value is not corrected for multiple testing of 10275 GO terms. "FDR q-value" is the correction of the above p-value for multiple testing using the Benjamini and Hochberg (1995) method. Namely, for the ith term (ranked according to p-value) the FDR q-value is (p-value * number of GO terms) / i. "Enrichment" is defined as follows: Enrichment = (b/n) / (B/N), where N - is the total number of genes; B - is the total number of genes associated with a specific GO term; n - is the number of genes in the top of the user's input list or in the target set when appropriate; b - is the number of genes in the intersection. "Gene examples" – certain genes with changed expression from the given GO Term.



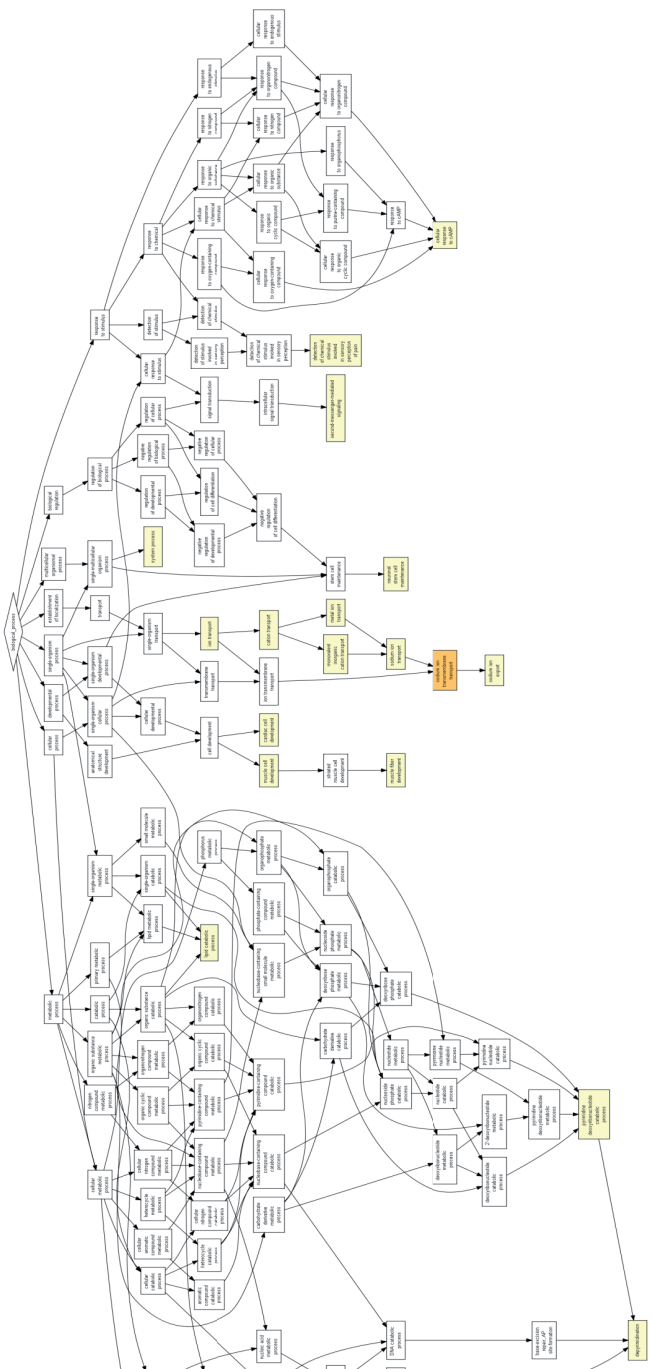
Supplementary Figure 1. Processes, affected by low salinity at 1 hour after the start of the experiment. Up-regulated genes. All the genes with up-regulated expression in worms at low salinity in comparison with starved worms (both at 1 hour after the start of the experiment) were put in groups based on their function (according to their GO-terms). All the processes these genes are responsible for are presented as boxes. The less transparent the box is, the more statistically significant is the enrichment for the particular GO term. It is mainly ion transport and general metabolic processes that are affected the most.



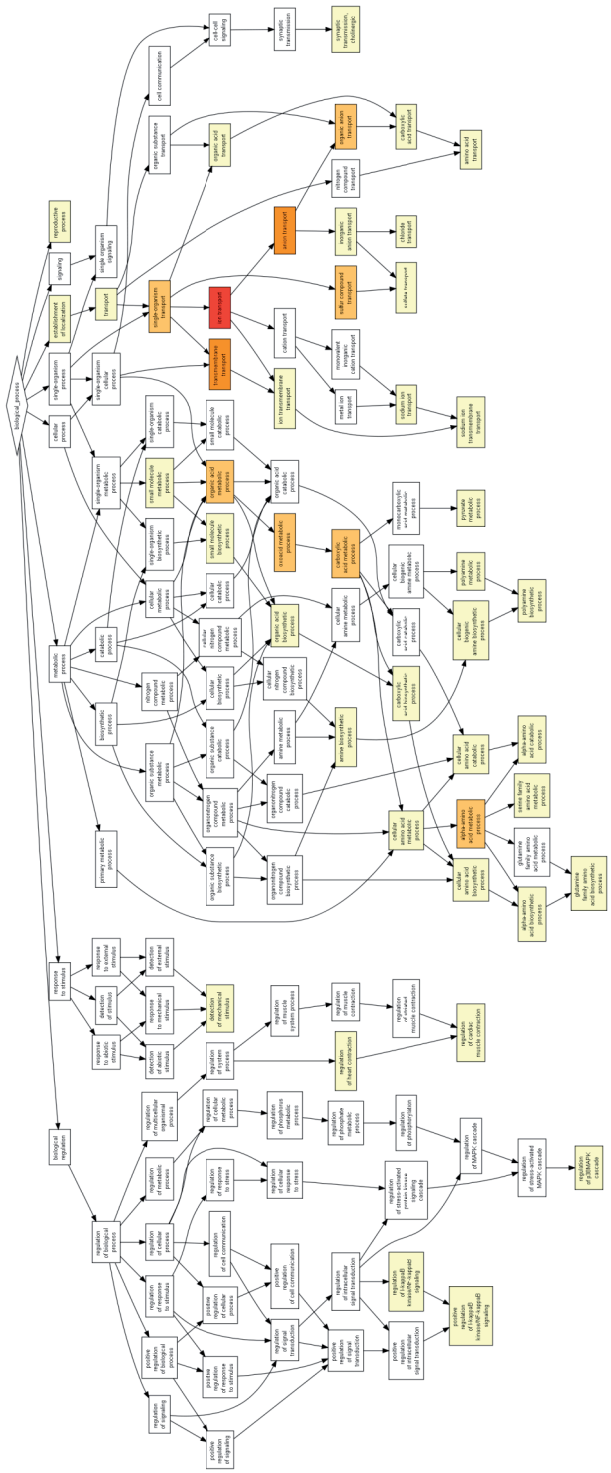
Supplementary Figure 2. Processes, affected by low salinity at 1 hour after the start of the experiment. Down-regulated genes. Here only the genes down-regulated in "low salinity" worms were used to construct the scheme. Apart from the transporters there is a big group of signal transduction genes (including the ones involved in stress-response signalling) present here.



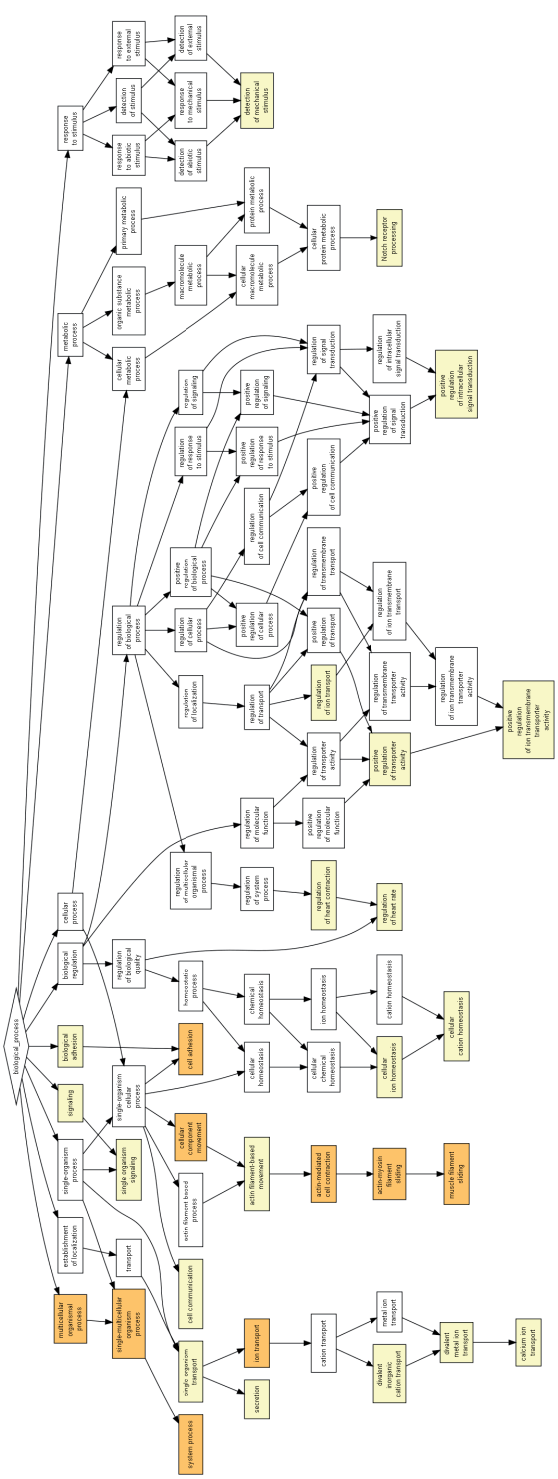
Supplementary Figure 3. Processes, affected by 60 hours of starvation. Up-regulated genes. Genes, that are up-regulated the most in starved worms at 60 hours after the beginning of the experiment are the ones responsible for metabolic processes, endosomal transport and axon ensheathment, though functions of the genes , presumably involved in the last two processes are not clearly known.



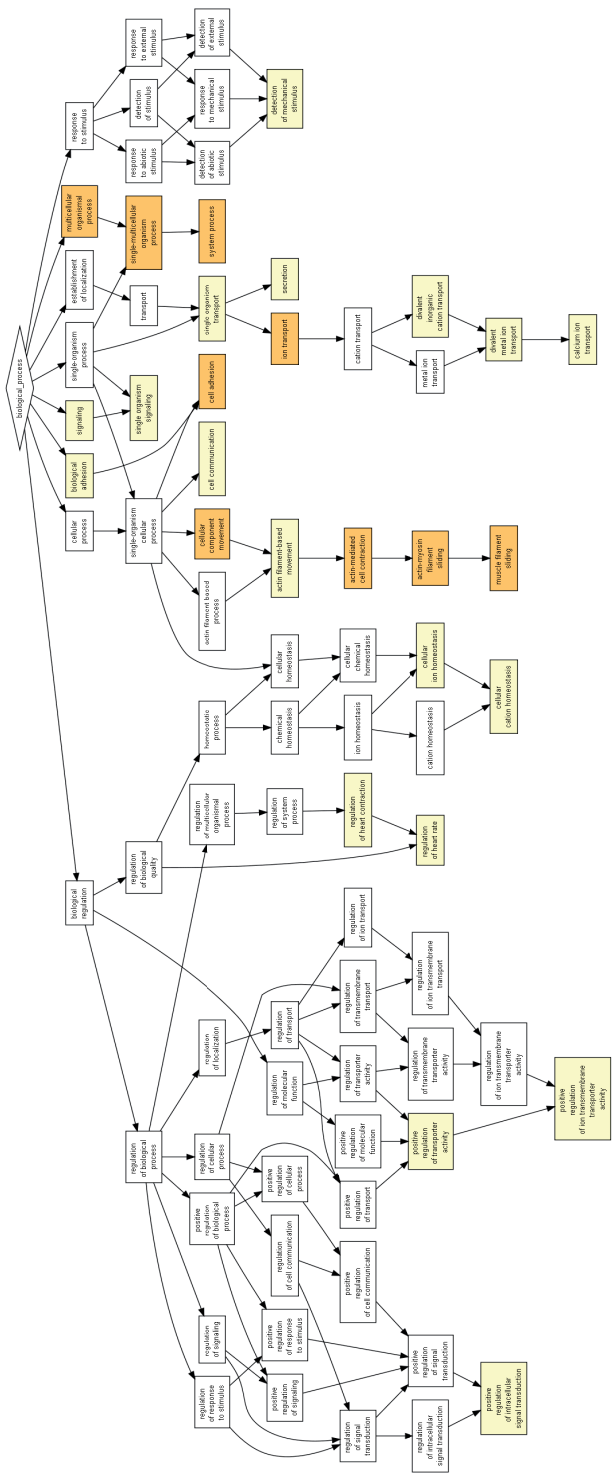
Supplementary Figure 4. Processes, affected by 60 hours of starvation. Down-regulated genes. Vast majority of genes, down-regulated by starvation is involved in various metabolic processes, such as lipid metabolism or nucleotide synthesis. However there are also genes involved in sodium ion trans-membrane transport and some signaling pathways members.



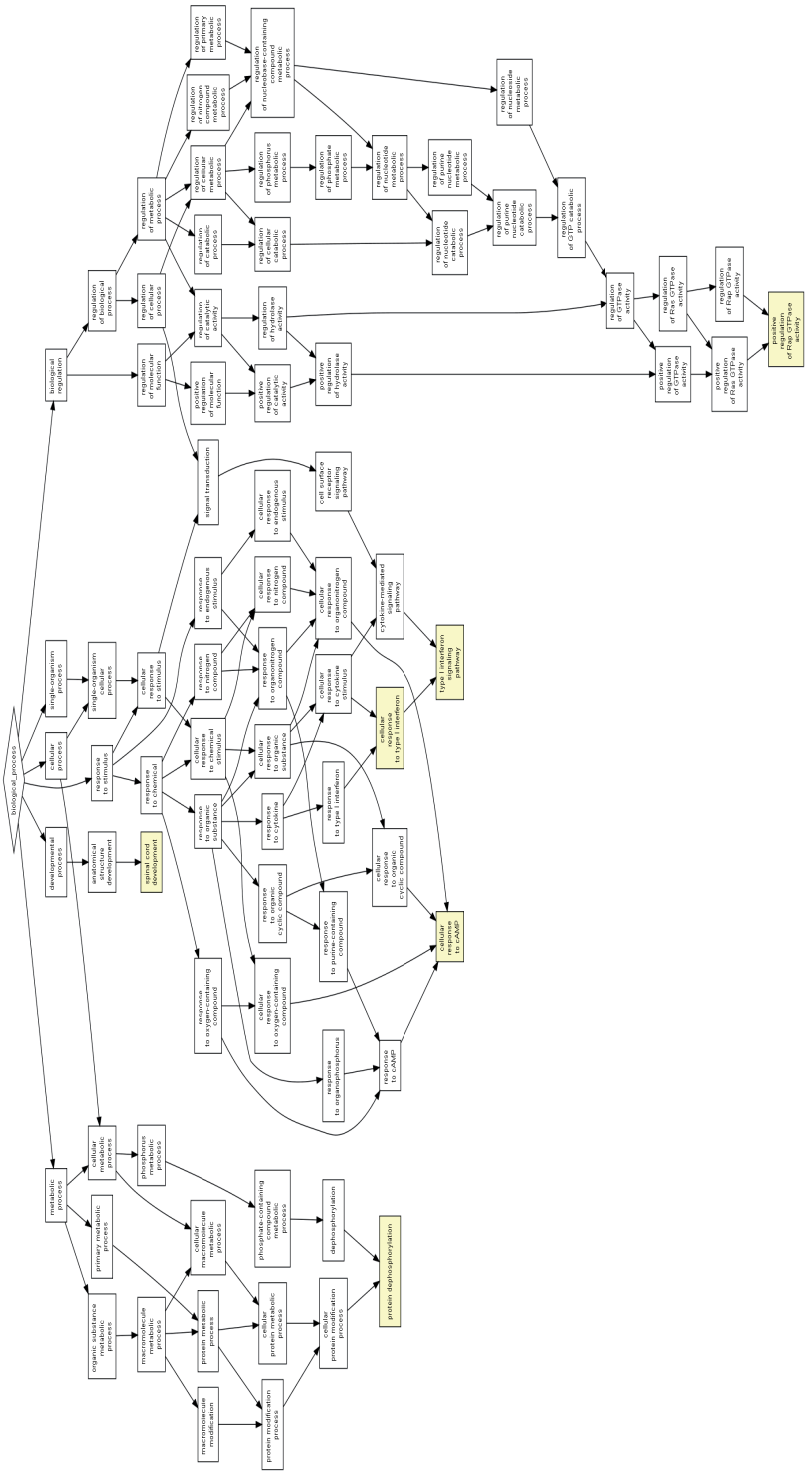
Supplementary Figure 5. Processes, affected at 60 hours of low salinity conditions in comparison with 60 hours of starvation. Down-regulated genes. Most genes, down-regulated in "low salinity" worms in comparison with starved worms at 60 hours after the start of the experiment are involved in ion transport and transmembrane transport or in metabolic processes. There is a large group of genes here, however, playing role in gonad functioning, and also a few apoptosis-related genes.



Supplementary Figure 6. Processes, controlled by genes up-regulated at low salinity conditions 60 hours after the change of the medium in comparison with starved worms at the same time point. Genes up-regulated in "low salinity" worms 60 hours after the change of the medium are involved in ion transport, but also in cytoskeleton construction and rearrangement, and formation of cell-to-cell contacts. There is also a small group of transcripts involved in signaling pathways.



Supplementary Figure 7. Processes, controlled by genes down-regulated at the first hour in low salinity conditions, and not afterwards. A lot of down-regulated genes are involved in ion transport. Processes, associated with cell adhesion and cell compartment movement are also significantly affected.



Supplementary Figure 8. Processes, controlled by genes up-regulated at the first hour in low salinity conditions, and not afterwards. Processes, affected the most are associated with camp and type I interferon signaling pathways.

CHAPTER 6

MIRNAS OF FLATWORM *MACROSTOMUM LIGNANO*: BIOGENESIS, EXPRESSION PROFILING AND POSSIBLE ROLE IN STEM CELL FUNCTIONING

Daniil Simanov[1], Katrien de Mulder[1], Willemijn Gommans[2], and Eugene Berezikov[1],[3].

[1]Hubrecht Institute, Hubrecht Institute, KNAW and University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands
[2]InteRNA Technologies B.V. Padualaan 8, 3584CH Utrecht, The Netherlands
[3]European Research Institute for the Biology of Ageing, University of Groningen, University Medical Centre Groningen, 9713AV Groningen, The Netherlands

Manuscript in preparation.

ABSTRACT

MicroRNAs (miRNAs) facilitate fine-tuning of tissue- and cell-specific gene expression, playing a crucial role in regulation of different biological processes, including stem cell self-renewal and differentiation. The role of specific miRNA families in proper functioning of different stem cell populations was recently shown *in vitro*. However, we still know close to nothing about stem cell specific expression of miRNAs and their functions *in vivo*. Flatworm *Macrostomum lignano* is a new attractive model for stem cell research famous for its impressive regeneration capacity, facilitated by a population of adult stem cells, known as neoblasts. Here we report the first large-scale sequencing and analysis of miRNAs in *M. lignano* and describe 185 miRNA genes, 118 of which identifiable homologs of known miRNAs from other species. Comparison of sequencing data from irradiated animals (in which stem cells were completely ablated) and control worms revealed 58 miRNAs potentially enriched in neoblasts and, thus, involved in stem cell functioning. We also describe highly unusual structure of miRNA hairpins, characteristic for *M. lignano* and discuss its possible evolutionary importance. Finally, we present the first results of miRNA *in situ* hybridization in *M. lignano*, as an easy and robust method to determine expression patterns of small RNA molecules.

INTRODUCTION

In recent years small non-coding RNAs have emerged as important players in virtually all biological processes, and microRNAs (miRNAs) together with Piwi-interacting RNAs (piRNAs) and small interfering RNAs (siRNAs) represent at the moment the most well-studied groups of small non-coding RNAs. Other small non-coding RNA species have been described in different model organisms, but their biological functions and exact mechanisms these RNAs are parts of remain largely unclear (Ghildiyal and Zamore, 2009; Aalto and Pasquinelli, 2012; Jiao and Slack, 2013). MiRNAs are single-stranded short RNA molecules that regulate gene expression by modulating mRNA activity. MiRNAs bind (in an partially complementary way in case of animals) to 3' untranslated region (3'-UTR) of specific target mRNA and either inhibit its translation or induce degradation of the transcript (Ambros, 2004; Bartel, 2004). Importantly, one miRNA can target a large number of mRNAs, and/or many miRNAs can bind to one specific mRNA (Filipowicz et al., 2008).

In the genome a miRNA is encoded as a continuous sequence of hundreds or even thousands base pairs and is transcribed as a capped and poly-adenylated long primary microRNA that contains a hairpin structure (Cai et al., 2004; Lee et al., 2004). This primary transcript is processed by RNAse III enzyme Drosha or alternative pathways (Fig. 1a-e) into a precursor miRNA (pre-miRNA), a stable hairpin of 80-120 nucleotides in length (Morlando et al., 2008). Structure of this hairpin is important not only for further processing and maturation of functional miRNA, but also for the evolution of miRNAs, since a very small change in the pre-miRNA hairpin structure can potentially lead to the generation of new miRNA species through the processes known as arm switching (Liu et al., 2008; Griffiths-Jones et al., 2011) and hairpin shifting (de Wit et al., 2009). Pre-miRNA is further processed by another RNAse III called Dicer, into 22 nucleotides long duplex (Fig.1f). This duplex is then targeted to miRNA-specific Argonaute protein that selects one of the strands of the duplex as a main one and brings it to target mRNA (Schwarz et al., 2003; Förstemann et al., 2007; Tomari et al., 2007; Tran and Hutvagner, 2013).

In mammals, miRNAs are predicted to control the activity of approximately one third of all protein-coding genes, and have been shown to participate in the regulation of almost every cellular process investigated so far (Filipowicz et al., 2008). This includes, for example, development (Ambros, 2003), stress response (Ambros, 2003; Bhattacharyya et al., 2006), fine-tuning of different physiological pathways (Ambros, 2003; Karres et al., 2007) and stem cell functioning – cell fate switches, proliferation and differentiation. Interestingly, the first two miRNAs discovered, *lin-4* and *let-7*, were characterized during the developmental stage transition in *C. elegans* (Lee et al., 1993; Reinhart et al., 2000). Later studies that involved deleting miRNA-processing enzymes showed that miRNAs are important in maintaining embryonic stem cell pluripotency and differentiation capacity (Murchison et al., 2005; Wang et al., 2007; Berardi et al., 2012). MiRNAs also play a role in the differentiation and self-renewal of mesenchymal stem cells (MSCs) (Guo et al., 2011). Below we describe a few specific examples of miRNA-mediated regulation of stem cell functioning.

One of the defining properties of stem cells is their ability to self-renew (Macarthur et al., 2009), and transcription factors *Oct4*, *Sox2*, and *Nanog* are known to play crucial role in this

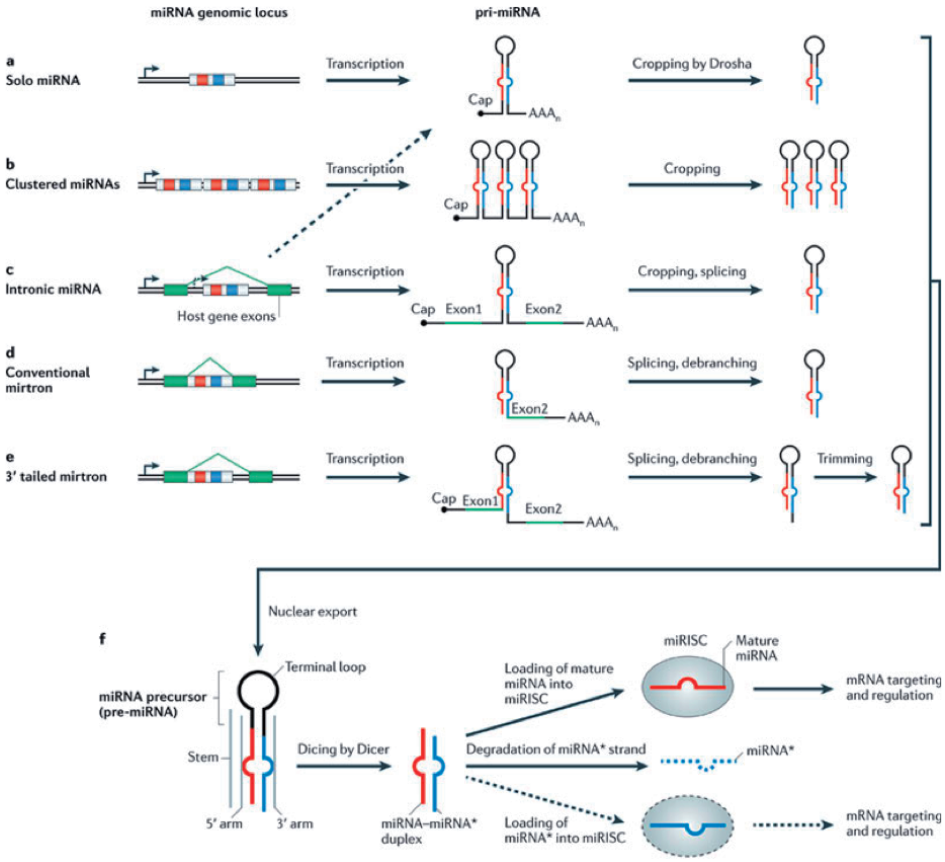


Figure 1. miRNA biogenesis. The distinguishing feature of microRNA (miRNA) genes is the folding of their intermediate RNA transcripts into hairpin structures that are specifically recognized and processed by the miRNA biogenesis machinery. **(a)** miRNAs are encoded in genomes either as independent transcriptional units with their own promoters (solo miRNAs) or **(b)** as clusters of several miRNA genes transcribed together. **(c)** A substantial fraction of animal miRNA genes are located in introns of protein-coding genes (Rodriguez et al., 2004). For all miRNAs, the hairpin regions in pri-miRNAs (also called stem–loop structures) are recognized and processed by the microprocessor complex, containing endonuclease Drosha. The result of the cleavage is a precursor miRNA (pre-miRNA). **(d)** Some unconventional intron-encoded miRNAs (called mirtrons), use the splicing machinery instead of Drosha to generate pre-miRNAs (Berezikov et al., 2007; Ruby et al., 2007). **(e)** In some cases, splicing can result in tailed mirtrons, which require additional trimming by the exosome to produce a functional pre-miRNA (Flyntet al., 2010). **(f)** Processed pre-miRNAs are exported into the cytoplasm, where they are recognized and enzyme, Dicer. The dicing of pre-miRNA cleaves off the loop region of the hairpin and results in a ~22–23 nt long RNA duplex. The mature miRNA strand of this duplex is next loaded into an Argonaute-containing miRNA-induced silencing complex (miRISC), that recognizes target mRNAs by partial base-pairing. Adapted from (Berezikov, 2011)

process (Chambers et al., 2003; Masui et al., 2007; Pan and Thomson, 2007; Yu et al., 2007; Silva and Smith, 2008; Mullen et al., 2011). These factors can regulate *mir-302-367* cluster of miRNA, which is differentially expressed in human embryonic stem cells. MiR-302-367, in turn, is required for *Oct4*, *Sox2*, and *Nanog* expression (Suh et al., 2004; Barroso-delJesus et al., 2008; Marson et al., 2008). *Mir-290-295* cluster is highly expressed in mouse embryonic stem cells, is regulated by Oct4, and binds Oct4, Sox2, Nanog. Expression of the miR-290 family promotes the G1/S transition, which enables rapid proliferation (Houbaviy et al., 2003; Wang et al., 2008). Another important gene in stem cell state maintenance is *c-Myc* (Takahashi and Yamanaka, 2006), which is inhibited by let-7 miRNA. *C-Myc* also binds to the promoters of *mir-141*, *mir-200* and *mir-429* - miRNAs that inhibit differentiation in mouse embryonic stem cells (Lin et al., 2009).

MiRNAs are not only involved in maintaining the pluripotency and proliferation rate of stem cells, but also can affect differentiation. Studies of the roles of some miRNAs in differentiation of vascular, endothelial, smooth muscle cells and cardiomyocytes are summarized in a few recent reviews (Howard et al., 2011; Ohtani and Dimmeler, 2011; Heinrich and Dimmeler, 2012; Choi et al., 2013). Expression of several miRNAs (let-7b, let-7f, miR-126, miR-130a, miR-133a, miR-133b, miR-210, and miR-296) is shown to be associated with angiogenesis and is normally upregulated in day 10 differentiated cells compared to embryonic stem cells (Kane et al., 2010). The miRNAs miR-143 and miR-145 are abundantly expressed in smooth muscle tissue and promote its differentiation from neural crest stem cells (Cordes et al., 2009). Cardiomyocyte differentiation is regulated by miR-1 and miR-133 that play opposing functions (Kwon et al., 2005; Zhao et al., 2007; Ivey et al., 2008). The same two players are also involved in skeletal muscle differentiation (Chen et al., 2006). Additionally, miR-204 promotes (Xiao et al., 2012) and miR-124 inhibits the cardiomyocyte differentiation (Cai et al., 2012). Expression of some miRNAs, such as the let-7 family, miR-124, and miR-9 is specific for neural tissue, and is essential for neurogenesis (Kawahara et al., 2012; Lang and Shi, 2012). For example, brain-specific miR-124 is upregulated during CNS development (Papagiannakopoulos and Kosik, 2009). Additionally, miR-137 and miR-184 regulate the balance of neural stem cell proliferation and differentiation in adult neurogenesis (Szulwach et al., 2010; Liu et al., 2010).

miR-29 has multiple distinct functions at different stages of osteoblast differentiation, controlling the cell fate all the way through, both directly and through the cross-talk with canonical Wnt pathway (Li et al., 2009; Kapinas et al., 2009; Lian et al., 2012). A few miRNAs contribute to osteoclast differentiation. The most well-known among them are miR-223 (Sugatani and Hruska, 2007) and miR-155, which can also serve as a key regulator in the maturation of hematopoietic cells to macrophages (O'Connell et al., 2007). Some miRNAs, such as miR-150, are involved exclusively in blood cell differentiation (Xiao et al., 2007).

Most of the studies briefly mentioned above were carried out *in vitro*. Specifically, while the role of miRNAs in differentiation of certain multipotent stem cell populations was in some cases addressed in different model organisms (such as mouse, zebrafish and flies), the effect of these small RNAs on the pluripotent stem cells was almost exclusively studied in cell cultures. There is a great need for the convenient model to explore the role of miRNAs in pluripotent stem cells *in vivo*, looking also at the niches of the cells. Flatworms can perfectly serve as such model, and

recently a number of papers was published, describing miRNAs in planarian flatworms and their potential roles in stem cell functioning (Resch and Palakodeti, 2012).

Planarian flatworms are small, grow fast and are easy to culture in the laboratory conditions (Sánchez Alvarado, 2004; Gentile et al., 2011). They are famous for their impressive regeneration potential, facilitated by adult pluripotent stem cells, also called neoblasts that make them an attractive model for stem cell research (Reddien and Sánchez Alvarado, 2004; Agata et al., 2006; Rink, 2013). Neoblasts of planarian flatworms are accessible to manipulations within their microenvironment and thus can be studied *in vivo* (Sánchez Alvarado, 2007). Information on planarian neoblasts, their physiological properties, differential potential and heterogeneity is summarized in a few recent reviews (Sánchez Alvarado, 2007; Aboobaker, 2011; Baguñà, 2012; Rink, 2013). It is worth mentioning that single cell transplantation experiments showed that some planarian neoblasts could be truly totipotent (Wagner et al., 2011). Molecular toolkit developed during last 15 years in this model system is quite impressive and contains sequenced genome and transcriptome, number of specific antibodies, established cell sorting, *in situ* hybridization (ISH) and RNA interference (RNAi) protocols and many other features (Umesono et al., 1997; Sánchez Alvarado and Newmark, 1999; Newmark and Sánchez Alvarado, 2000; Newmark et al., 2003; Hayashi et al., 2006; Robb et al., 2008; Pearson et al., 2009).

The first description of miRNAs expressed in planarian flatworm *Schmidtea mediterranea* was made in 2006 (Palakodeti et al., 2006), and the first successful miRNA expression pattern analysis by *in situ* hybridization was performed in 2009 (González-Estévez et al., 2009). A number of miRNAs was identified in several planarian and also parasitic flatworm species, and many of these were highly conserved across the animal kingdom (Lu et al., 2009; Huang et al., 2009; Simões et al., 2011; Qin et al., 2012; Xu et al., 2013). Deep sequencing of small RNA libraries from irradiated and non-irradiated animals, sexual and asexual strains and sorted stem cell populations revealed some examples of strain-specific and neoblast-specific miRNA expression in *Schmidtea mediterranea* (Lu et al., 2009; Friedländer et al., 2009; Resch et al., 2012; Sasidharan et al., 2013). These experiments revealed that miRNAs belonging to the major miRNA cluster *mir71a/2d/13/752* are sensitive to irradiation and are likely to be expressed in the neoblast population, suggesting also that miRNAs belonging to the same cluster may be co–regulated (Lu et al., 2009; Friedländer et al., 2009; Sasidharan et al., 2013). Interestingly, *let-7a* miRNA expression is decreased in planarias after irradiation, while *lin-28,* a suppressor of *pre-let-7a* processing, was expressed in differentiated tissues. In contrast, in mammalian embryonic stem cells *let-7a* expression is suppressed, and *lin-28* is upregulated (Lu et al., 2009). Important genes that encode proteins essential for miRNA biogenesis were identified in planarian genome. Knock downs of some of these genes in *Dugesia japonica* and *Schmidtea mediterranea* leads to defects in regeneration, suggesting that planarian miRNAs are important for regeneration and neoblast function *(Rouhana et al., 2010; Li et al., 2011)*.

With the help of flatworm models the first step towards understanding the role of miRNAs in pluripotent stem cell regulation has been made, but we still know almost nothing about expression and function of specific miRNAs. Additionally, miRNA studies in flatworms were so far limited to planarian species that have certain limitations as model systems. The biggest disadvantages of planarian flatworms are, first of all, lack of transgenics and, secondly, difficulties

with utilizing forward genetic approaches due to the struggles with sexual reproduction and ability to reproduce asexually, characteristic for these species.

Our group uses another flatworm species, *Macrostomum lignano*, as a model organism for studying stem cells. It is a small seawater free-living flatworm that represents a basal member of free-living Platyhelminthes (Ladurner et al., 2005). *M. lignano* has recently proven itself as a useful invertebrate model for stem cell, regeneration, aging and sex allocation research (Egger et al., 2006; Pfister et al., 2008; De Mulder et al., 2009; Mouton et al., 2009; Verdoodt et al., 2012; Simanov et al., 2012; Janicke et al., 2013; Demircan and Berezikov, 2013), and knowing more about miRNAs and their role in this organism would be important for better understanding of the regulation of all these processes. There is an extensive molecular toolkit available for *M. lignano* that includes ISH (Pfister et al., 2007) and RNAi (Pfister et al., 2008; Demircan and Berezikov, 2013) techniques, immunohistochemistry (Nimeth et al., 2004; Ladurner et al., 2005; Pfister et al., 2007), set of tissue-specific antibodies (Ladurner et al., 2005) and sequenced and annotated draft genome and transcriptome (see http://www.macgenome.org and Chapters 2 and 3 in this Thesis). Most importantly, *M. lignano* is an obligatory non-self-fertilizing hermaphrodite and has exclusively sexual reproduction. Worms lay a lot of eggs all year through making them amenable for genetic manipulations and providing the material for generation of transgenics. The first transgenic lines, including the ones generated by site-specific mutagenesis, are already available (Demircan et al, in preparation). Access to transgenics makes it possible to create worms with miRNA knockouts (and conditional knockouts) or animals overexpressing certain miRNAs. Thus, transgenics would provide a great tool to study miRNAs in pluripotent stem cells *in vivo* in *M. lignano* in great detail.

Here we present the results of deep sequencing of small RNA libraries from *M. lignano*. We identified 185 distinct miRNAs expressed in this model organism. Most of them are conserved with other species, while there are also *Macrostomum*-specific miRNAs. We describe the unusual structure of pre-miRNA hairpins, characteristic for many of *Macrostomum* miRNAs, which might be a hallmark of their great evolutionary age. We also compared small RNA deep sequencing data from irradiated (neoblast-ablated) and non-irradiated worms to identify miRNAs upregulated in stem cells and potentially involved in neoblast regulation. Finally, we present the first results of mature miRNA in situ hybridization (ISH) utilizing double DIG labeled miRcury LNA probes as a robust method to identify expression patterns of different miRNAs in *M. lignano*.

RESULTS AND DISCUSSION

1. Deep sequencing of *M. lignano* microRNAs

To identify miRNA species involved in stem cell regulation, we wanted to profile expression differences of miRNAs between irradiated (neoblast-ablated) and control worms. Therefore, small RNA fraction was obtained from two populations of *M. lignano* – normal worms and animals irradiated with repetitive high doses of γ-irradiation (De Mulder et al., 2010). Both samples were sequenced with Illumina platform and obtained reads were mapped to the genome assembly ML100925 (See Chapter 2 and http://www.macgenome.org). In total,

10,935,401 sRNA reads were generated (5,9 million in control library and 5 in the one prepared from irradiated worms), of which 8,356,389 were of miRNA size range (18-28 nucleotides) with identifiable cloning adapter at 3' end. 7,056,460 reads were mapped to the genome assembly ML100925 with perfect matching and used in the subsequent analysis for identification of miRNA genes. Information about the reads generated, accepted and mapped and about the size distribution of the mapped reads in both libraries is summarized in the Figure 2. Generally, quality of both libraries is very good and libraries are comparable with each other. A lot of good quality reads of miRNA size were generated in both libraries (82% and 70% in control and

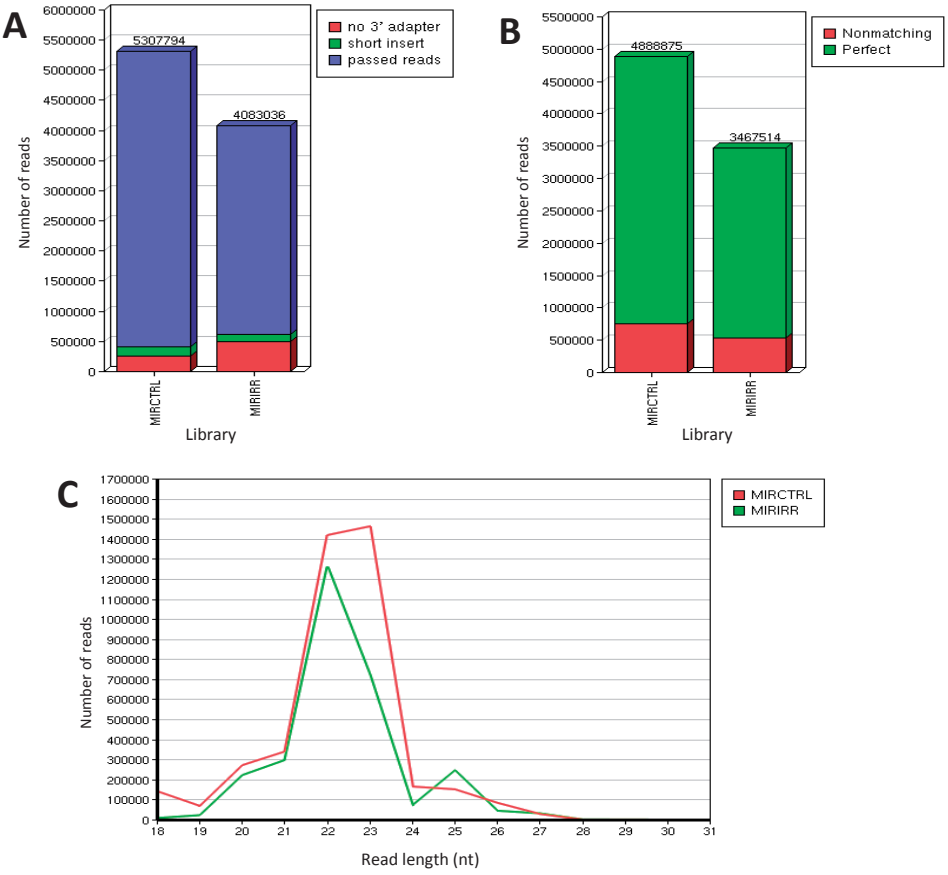


Figure 2. Read quality and genome mapping data for both small RNA libraries. MIRCTRL – library prepared from the control worms; MIRIRR – from irradiated, neoblast-ablated animals. **(A)** Read quality and adapter processing. From the control library almost 1 million reads more were generated, than from irradiated worms. 82% of the control reads and 70% of the reads from the "irradiated" library pass as possible miRNA reads. In the control library 10,5% and in irradiated one 18,2% of reads failed being longer than expected miRNA size (not shown). 2,3% and 1,1% of reads respectively were too short, and 4,3 and 9,9% did not have 3' adapter. Around 1% of the reads in both libraries did not have an insert at all. **(B)** Genome mapping. 84% of all passed reads in both libraries mapped perfectly to the genome. **(C)** Length distribution of mapped reads, with the peak at 22 nt characteristic of miRNAs.

"irradiated" libraries respectively) (Fig. 2A), and the majority of them (84% in both libraries) mapped perfectly to the genome (Fig. 2B).

2. Identification of conserved miRNAs and discovery of novel miRNAs

To identify conserved and novel miRNAs from generated sRNA reads, we used previously developed miR-Intess discovery and annotation pipeline (Berezikov et al., 2006; Berezikov et al., 2010) with the parameters tuned to allow identification of longer than usual for miRNAs hairpins (up to 1 kb). This adjustment was necessary since initial search for miRNAs with parameters tuned for discovery of miRNAs in *C. elegans*, *Drosophila* and human failed to identify several broadly conserved miRNAs, such as *mir-124,* although sequence of mature miR-124 was present within *M. lignano* sRNA reads and fully conserved. Detailed investigation revealed that *mir-124* can be identified in *M. lignano* genome but its pre-miRNA formed a hairpin of 250 nt (see details below), a substantially large size than a typical for this pre-miRNA 120 nt. In total we identified 185 miRNA loci in *M. lignano* genome assembly ML100925 using the generated sRNA data. These loci are classified into miRBase homologs, novel and candidate miRNAs based on their alignment with known miRBase miRNAs (Griffiths-Jones et al., 2006; Kozomara and Griffiths-Jones, 2011) and characteristics of miRNA precursor stem-loop.

35% of all the mapped reads from the both libraries corresponded to these 185 miRNAs. All of them are listed in the Supplementary Table 1 and further details are provided at http://www.macgenome.org/pub/ML_140120.

From the 185 identified miRNA loci 118 miRNAs have homologs in miRBase are can be further divided into two subgroups based on the level of conservation of their pre-miRNA sequence and secondary structure. 22 miRBase-related miRNAs are conserved among different animal species and their pre-miRNA sequences align nicely to known pre-miRNAs from miRBase database and all the characteristics of their hairpins comply with the standards known for miRNAs (such as Drosha/Dicer signature, number and size of bulges and loops in hairpins), and alignment of their secondary structures with RNAforester (Höchsmann et al., 2004) shows high similarity. For these "obviously conserved" miRNAs the average distance between the arms of pre-miRNA and the terminal loop (as a sum of distances between each arm and the loop) is around 17-18 nucleotides. The other 96 identified miRBase-related miRNAs also possess all the characteristic features of miRNAs but fail to align to miRBase pre-miRNAs beyond the region of mature miRNA and do not have acceptable RNAforester score due to substantial length differences in the arm/loop regions. The same distance here is on average more than 40 nucleotides, and in some cases arms are separated from the terminal loop by more than 70 nucleotides in total. However, the conservation in the mature miRNA sequences allows to reliably assigning them as homologs of known miRNAs. This group of miRBase-related miRNAs, where alignment on pre-miRNA level is problematic, is exemplified by sblock19584. It is a clear homologue of *mir-124* (sequence of mature sblock19584 is exactly the same, as sequence of human miR-124), but it has an extremely long pre-miRNA sequence with the distance between arms of 111 nucleotides (Figure 3).

3% and 31% of all mapped sRNA reads (or 8% and 89% of all miRNA reads) from both libraries corresponded to these two subgroups of miRBase-related miRNAs respectively. The 118 identified miRBase miRNA homologs can be grouped into 49 distinct families based on the miRBase annotation of the homologs of the most abundant sRNA read (Sup. Table 1).

Next to miRBase-related miRNAs we identified only 11 novel miRNA, which pass all the stringent criteria for miRNA annotation (Sup. Table 1). Most of these novel miRNAs are expressed at low levels compared to miRBase-related miRNA (the total fraction of reads corresponding to novel miRNAs is 1%). At the same time, there are several novel miRNAs with high expression level, for example sblock20297 (Figure 4).

Finally, we identified a group of 56 miRNA candidates (Sup. Table 1). These are the loci that possess some of the characteristics of miRNAs but also have some atypical features (e.g. very low read count, large number of genomic loci, suboptimal hairpins, weak Drosha/Dicer signature) that prevent their reliable classification at this point and more sRNA data need to be generated in order to resolve their classification.

MiRNAs often form clusters of several genes, which are transcribed as single pri-miRNA. In our analysis we identified 21 such clusters consisting of 2 or 3 miRNAs (Sup. Table 2). However, it should be noted that due to the draft nature of the used genome assembly there might be some clusters missing or misannotated and detailed analysis of miRNA clusters will require improvements in genome assembly continuity.

3. Unique hairpin structure of *M. lignano* microRNA precursors

More than half of all identified miRNAs have unusual structure of their stem-loops. The length of the hairpin is in a lot of cases much bigger than "normal" 80-120 nt, and so is the distance between two arms. In a lot of cases there are also many bulges and some secondary hairpin structures on the loop. Until very recently increased length and number of bulges and internal loops were considered to be a distinctive feature of non-precursor RNA duplexes, that cannot be processed and give origin to mature miRNAs (Ritchie et al., 2007). But in the last few years

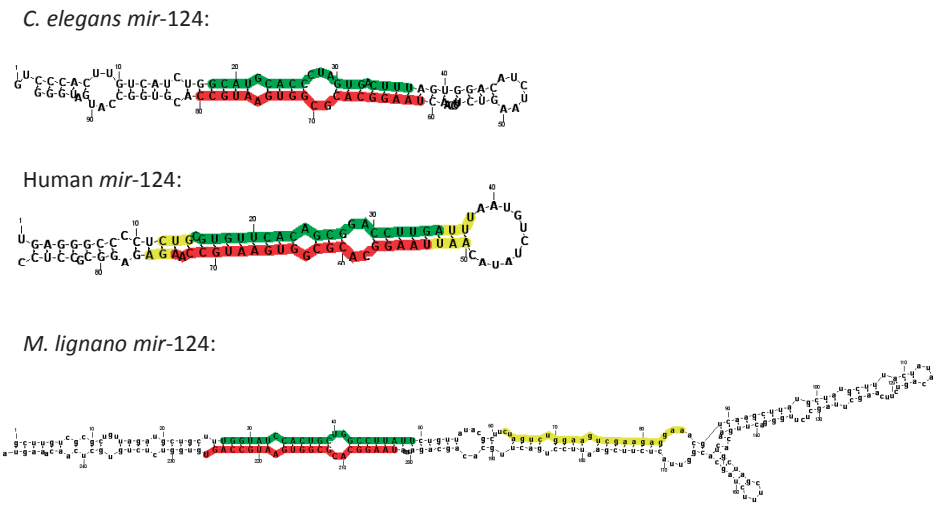


Figure 3. Stem-loops of *mir-124* precursors from different species. Mature miRNA sequence (red) is highly conserved, but stem-loop structure of *M. lignano mir-124* is more complex than of its *C. elegans* and human homologues with the distance between arms much longer.

a lot of miRNAs were identified in different species that do have precursors of unusual size and composition, and new popular algorithm for miRNA discovery do not take into account loop length and complexity of the structure as the most important characteristics of true pre-miRNA (Schwab and Voinnet, 2009; Song et al., 2013).

Still, the fact that 41% of miRNAs expressed in *M. lignano* have elongated stem-loop (longer than 120 nt) with unusually complex structure is surprising. Interestingly, the conformation of the hairpin does not seem to impair transcription, processing and maturation of miRNAs, since the most abundantly expressed miRNAs have stem-loops longer than 140 nt (Fig. 5A), and, generally, miRNAs with stem loop length from 80 to 170 nt can be expressed at high levels (more than 10000 reads per transcript, Fig. 5B).

Why do miRNAs of *M. lignano* have such an unusual complex structure of their stem-loops that is more characteristic for plants than for animals (Axtell et al., 2011) remains unclear. Some of *M. lignano* miRNAs with long hairpins, like sblock19584, are clearly conserved and homologous to miRNA from other species with more canonical stem-loop structure. Thus, it is highly unlikely that *Macrostomum* developed its own unique set of miRNAs with complex precursors and some alternative processing mechanism. According to one of the models of animal miRNA evolution, miRNA hairpins can emerge from the initially unstructured sequences in the non-coding parts of the genome. These small RNAs could potentially have any structure. Only once they acquire a certain target, they get under pressure of evolution and change

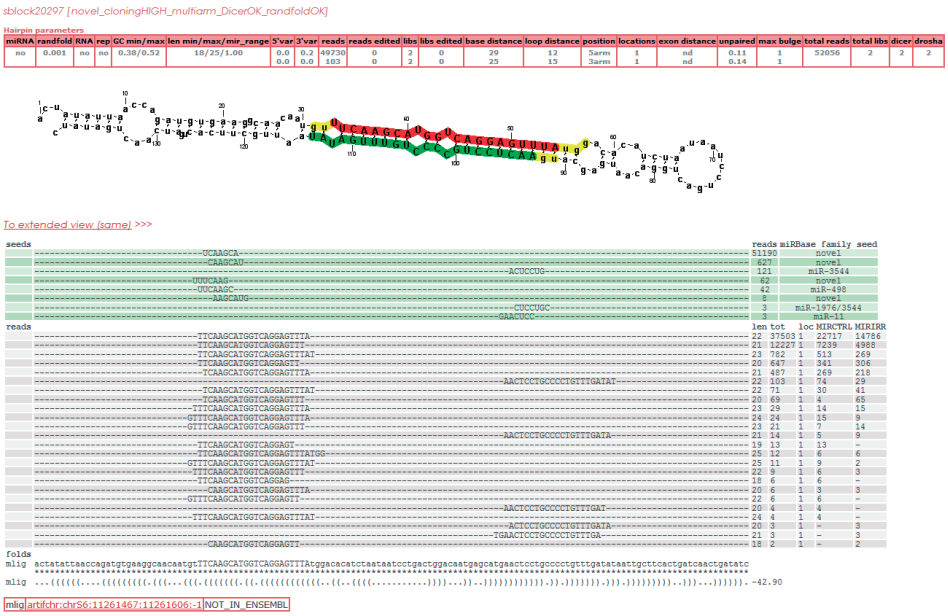


Figure 4. An example of novel miRNA in *M. lignano*. miR-Intess parameters used for hairpin classification are show at the top, followed by the hairpin structure, with mature and star regions show in red and green respectively. Read alignments and counts are shown below, grouped by the 8-mer "family seeds" (nucleotides 2-8 of mature miRNA). The most abundant reads do not correspond to any known family seed in miRBase.

Figure 5. Stem-loop size and expression of miRNAs. **(A)** Percentage of miRNAs with stem-loops of given size (in blue) and percentage of sRNA reads corresponding to the transcripts with the stem-loop of given size (in red). **(B)** Scatter plot showing expression of miRNAs (reads in both libraries) depending on their stem-loop size.

selectively towards the conformation that can be more efficiently processed by miRNA maturation machinery (Ruby et al., 2007; Lu et al., 2008; Berezikov et al., 2011). From this point of view, unusual length and structure of pre-miRNAs can be an indication of evolutionary old, basal set of miRNAs. Interestingly, in other basal species the same unusual miRNA stem-loops were observed. For example, in a sponge *Amphimedon queenslandica* pre-miRNA hairpins are reported to be larger than in most of other animals, while in sea anemone *Nematostella vectensis* miRNA stem-loops are significantly shorter than normal (Grimson et al., 2008).

Further studies of *M. lignano* miRNAs and protein machinery responsible for miRNA processing and functioning and its detailed comparison to miRNA system of cnidarians and sponges would provide an insight into miRNA evolution and can possibly help to reconstruct the ancient set and structure of miRNAs, characteristic for hypothetical metazoan common ancestor.

4. Candidate stem cell specific microRNAs

Comparison of miRNA expression levels in irradiated and control worms reveals a number of miRNA species significantly down- or up-regulated in neoblast-ablated animals (Tables 1 and 2). In some miRNA families it is impossible to distinguish the reads between different family members. In these cases we only discuss expression of one miRNA, keeping in mind that the rest of the family is most likely regulated and expressed in the same way. That leaves us with the list of 139 miRNAs expression dynamics of which we can compare. There are much more miRNAs, expression of which is decreased after irradiation, than the opposite. In total only 12 transcripts (expressed at reasonable level, so with at least 20 reads together in both libraries) show at least 1,5-fold increase in expression level after irradiation. All of them are shown in Table 2. At the same time 58 miRNAs (with the same level of basal expression) demonstrate at least 1,5-fold decrease in expression after irradiation. Top 20 down-regulated miRNAs are listed in the Table 1. Additionally, decrease in expression level after irradiation is more crucial than increase. 12 up-regulated miRNAs demonstrate on average 8-fold change, while mean expression decrease of 12 top down-regulated is 33-fold.

As expected, up- and down-regulated miRNAs belong to different families. There are two members of miR-1175, miR-200 and miR-2049 families among 20 top down-regulated in irradiated worms, and two up-regulated miRNAs represent miR-973 family, suggesting that generally expression of miRNAs within the same family is regulated in the same way.

With respect to the conserved miRNAs identified in *M. lignano*, our data are supported by the previous studies. For example, miR-71b and miR-2a are likely expressed in planarian neoblasts, and in irradiated planarians expression of these miRNAs is decreased (Friedländer et al., 2009). Both of them are also down-regulated after irradiation in *M. lignano* (sblock18341 and sblock18339), and miR-2a even appears among top 10 most significantly down-regulated (Table 1; Sup. Table 1). Expression of miR-124, specific for central nervous system (González-Estévez et al., 2009), is slightly reduced in planarian species following irradiation (Friedländer et al., 2009), and its *Macrostomum* homologue sblock19584 demonstrates the same behavior (Sup. Table 1). miR-133 was recently shown to be down-regulated during the fin regeneration in zebrafish (Thatcher and Patton, 2010). It can be involved in regulation of canonical Wnt pathway, that plays an important role in flatworm regeneration as well (Gurley et al., 2008; Petersen and Reddien, 2008; Almuedo-Castillo et al., 2012). In *Macrostomum* expression of miR-133 (sblock13128) is significantly down-regulated after irradiation (about 6-fold, Sup. Table 1), suggesting its expression in neoblasts.

Table 1. Top 20 miRNAs down-regulated after irradiation

miRNA ID	Down	Family
sblock18922 (miR-50)	complete	miR-190//
sblock10345 (miR-1175)	66	miR-1175
sblock19586 (miR-3177)	62	miR-3177
sblock18771	59	novel-3
sblock19588	41	miR-2160
sblock18341 (miR-2a)	29	miR-11//
sblock17468 (miR-236)	24	miR-236
sblock11336 (miR-8a)	23	miR-200//
sblock6458 (miR-8a)	23	miR-200//
sblock19133 (miR-1175)	22	miR-1175
sblock20606 (miR-49)	19	miR-285//
sblock20293 (miR-1993)	17	miR-1993//
sblock6462 (miR-993)	16	miR-100//
sblock20587 (miR-87)	13	miR-233//
sblock19316 (miR-92c)	12	n/a
sblock12166 (miR-87)	11	miR-4293
sblock17917 (miR-33b)	10	n/a
sblock19546 (miR-96a)	10	miR-1271//
sblock19380 (miR-277c)	10	miR-2049//
sblock17744 (miR-277)	8	miR-2049//

Table 2. Top 15 miRNAs up-regulated after irradiation

miRNA ID	Up	Family
sblock19666 (miR-7)	35	n/a
sblock19255 (miR-745b)	32	n/a
sblock20438 (miR-92)	9	miR-973
sblock9888	6	miR-181d//
sblock19814 (let-7b)	3	miR-1961//
sblock19434	2	n/a
sblock8706 (miR-996)	1.8	n/a
sblock6327 (miR-219)	1.8	miR-1493//
sblock19237 (miR-5706)	1.7	miR-1820//
sblock18471 (miR-44)	1.7	miR-247//
sblock12796 (miR-92)	1.7	miR-973
sblock7686 (miR-10b)	1.7	miR-10b//

"Down" and "Up" are the fold-changes of decrease or increase in expression respectively. "Family" is the miRNA family given transcript belongs to. miRNAs, expressed abundantly, are shown in green (more than 2000 reads in two libraries together); in yellow – transcripts with average level of expression (500-2000 reads); Lowly expressed (20-500 reads) miRNAs are shown in grey. More information in the text.

5. MicroRNA in situ hybridization in *M. lignano*

For the future studies of small RNAs in *M. lignano* it is important to establish a reliable technique to identify expression patterns of miRNAs. We tested mature miRNA in situ hybridization approach, making use of commercially available miRCURY LNA (locked nucleic acid) double DIG-labeled probes from Exiqon. For the test we used miR-7 and miR-184 - two miRNAs that are known to be expressed in central nervous system (Aboobaker et al., 2005; González-Estévez et al., 2009; Choudhury et al., 2013) and therefore should have a very distinct expression pattern. Additionally, both miRNAs can be involved in nervous system development and regeneration (Candiani et al., 2011; Liu et al., 2010). We designed two probes against miRNAs sblock13174 (clear homologue of miR-184) and sblock19666 (the closest homologue of miR-7 in *M. lignano*), and tested them in ISH approach. miR-7 (sblock19666) appeared to be expressed in the brain (Fig. 6). miR-184 is also expressed in the brain, but originally the signal we saw was barely detectable. Increasing probe concentration and time of signal development leads to much stronger signal in the brain. Besides, it also appears in the testes (Fig. 6). Although miR-184 was previously reported to be involved in spermatogenesis (Wu et al., 2011), we cannot be sure whether this testes signal is specific (for under the same conditions it also appears in the tail region in a pattern characteristic for unspecific ISH signal).

Thus, the method we used appears to work fine to detect *Macrostomum* miRNA expression patterns, though some optimization steps and additional tests would be needed to make the best of it. One of the most commonly used methods to identify miRNA expression patterns (and the most commonly used in planarian flatworms) requires use of the probe against pri-miRNAs rather than mature miRNAs (González-Estévez et al., 2009), and they do not always have the same expression pattern. So the big advantage of the method we used is that it allows working directly with mature sequences.

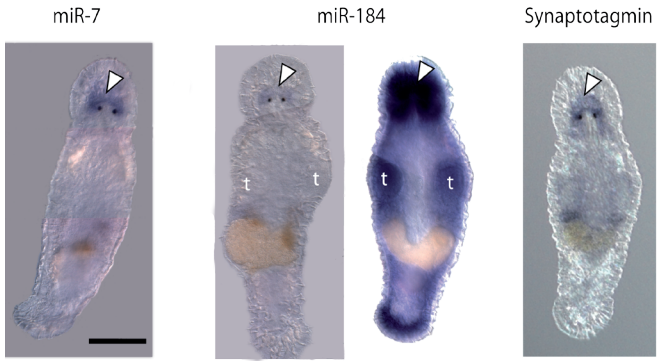


Figure 6. Expression patterns of *M. lignano* miR-7 and miR-184. (Left) miR- 7 (sblock19666) is expressed in the brain, as expected. MiR-184 (sblock13174) expression could also be detected in brain, and, when using higher probe concentration and longer development time, in testes and the tail plate region (we are not sure about specificity of testes and tail plate signal). Expression pattern of *Synaptotagmin,* detected by the standard ISH protocol is demonstrated as a canonical example of brain-specific signal. White arrowheads indicate the brain region, "t" are testes. Scalebar 200 μm (all worms).

CONCLUSION

We performed the first deep sequencing of *M. lignano* miRNAs. 185 miRNA sequences were identified, of which 27 are obviously conserved with other species annotated in miRBase, and 91 are conserved at the level of mature miRNA sequence but have differently sized stem-loop structure. Conservation of *M. lignano* miRNAs is challenging to address mainly because of the unusual and complex structure of pre-miRNA hairpin (namely, extended length, number of bulges and distance between the arms, and presence of additional loops) that appears to be a characteristic feature of most of *M. lignano* miRNAs. We consider this structure to be a hallmark of the evolutionary old miRNA system. Comparative analysis of miRNAs from *Macrostomum* and other basal species would provide a great insight into early evolution of animal miRNAs. Based on the comparison of expression data from irradiated and control animals we identified miRNAs that are potentially enriched in *M. lignano* neoblasts and could be involved in stem cell functioning. Finally, we show that double-DIG labeled LNA probes can be successfully used to identify mature miRNA expression patterns in *M. lignano*.

MATERIALS AND METHODS

1. Animal culture

M. lignano worms of inbred DV1 line (Janicke et al., 2013) were cultured in Petri dishes in f/2 medium, nutrient-enriched artificial sea water (Andersen et al., 2005), 150-200 animals per plate. Worms were fed with unicellular diatom *Nitzschia curvilineata*, which covers the bottom of the petri dishes (Ladurner et al., 2005). All animals were kept in the incubators under following conditions: 20°C temperature, 60% humidity and 14/10 h day/night cycle (Rieger et al., 1988).

2. Irradiation by γ-rays

Mixed populations of worms were used. Worms were irradiated by three single successive doses (210 Gy, 15 Gy and 15 Gy) in accordance with "fractionated irradiation" protocol described before (De Mulder et al., 2010). Control (mock-irradiated) animals were handled in the same way as treated worms, only irradiation steps were omitted. Both irradiated and control worms were sacrificed for total RNA extraction 7 days after the last irradiation step.

3. Total RNA extraction

Total RNA was extracted using TRI Reagent (T9424, Sigma), according to manufacturer's instructions. Animals were homogenized in TRI Reagent by pipetting. For every extraction a batch of 150-200 worms was used (all the worms from the plate). Samples were resuspended in nuclease-free water and treated with 5 U of DNAse I (Thermo Scientific, Cat. No. EN0521) for 45 minutes at 37°C. Enzyme and all the remaining DNA were removed by extraction with phenol: chloroform : isoamyl alchohol (125:24:1, pH 4,5 Life technologies, Cat. No. AM9720). Samples were alcohol precipitated overnight at −80°C. Total RNA was pelleted by centrifugation at 12,000g for 20 mins at 4°C, washed with 70% ethanol and air-dried for 5 minutes. RNA was resuspended in nuclease-free water. Concentration of total RNA samples was measured with Qubit RNA BR assay kit (Invitrogen, Cat. No. Q10211).

4. Preparation and sequencing of RNA-seq libraries

Small RNA fraction was separated from total RNA on 15% acrylamide/bis gel (100V, 2hours). RNA was extracted from gel piece with 0,3 M NaCl (overnight rotation at 4°C), and then precipitated with glycoBlue (Life Technologies, Cat. No. AM9516) and equal volume of isopropanol. RNA 5'- and 3'-adaptors were ligated to the sample subsequently; After each ligation fraction of sRNA with ligated adaptors (between 40 and 60 nt after the first reaction and around 70 nt – after second ligation) was separated from adaptor-dimers and non-ligated fragments on acrylamide gel and extracted. cDNA was prepared from adaptor-flanked sRNA sample, then the library was amplified by PCR. Reaction was precipitated with 100% ethanol (2,5x volume), 3M NaOAc (0.1x volume) and Glycoblue and dissolved in nuclease-free water. Then the library was run on 8% acrylamide/bis gel, band around 90 nt was cut, and DNA was extracted with NEBuffer 2 (NEB, Cat. No. B7002S) (overnight elution at 4°C) and Spin-X filter (Sigma-Aldrich, Cat. No. CLS8160). Flow-through was ethanol-precipitated once again and resuspended in 10 mM Tris–HCl Buffer, pH 8.5. Detailed protocol for library preparation is described in (Gommans and Berezikov, 2012).

sRNA libraries were then sequenced on Illumina platform as a service from Baseclear BV (Leiden, The Netherlands).

5. sRNA read analysis and identification of miRNAs

To process generated small RNA reads and identify miRNAs, we used previously developed miR-Intess discovery and annotation pipeline (Berezikov et. Al, 2006; Berezikov et al., 2010). Raw reads were first processed with custom perl scripts to trim adapter sequences, remove short reads and collapse redundancy. Next, sRNA reads were mapped to ML100925 genome assembly using BLAST (Altschul et al., 1997) and only perfect hits were retained for further analysis. Overlapping read matches were grouped into blocks by genomic coordinates and these blocks were used as anchors to predict potential hairpin structures around them using RNAfold software (Lorenz et al., 2011) and allowing for hairpins of up to 1 kb in length. Predicted overlapping hairpins from adjacent blocks were merged thus allowing calculation of Drosha/Dicer signatures and identification of mature and star reads.

For the classification of hairpins into confident miRNAs, candidate miRNAs and other hairpins a number of parameters were calculated, including the stem-loop size and the number of wobbles, Drosha/Dicer signature, randfold value (Bonnet et al., 2004), GC content, number of locations in the genome and cloning frequencies. Preliminary automated classification of hairpins was performed using these parameters and thresholds tuned to identify the majority of known miRNAs (based on data from known organisms) while keeping the number of other hairpins in the same 'confident' category to the minimum. The automatic classification was manually curated to make final assignments of hairpins to a particular category.

For the analysis of hairpin conservation in other genomes, predicted hairpins sequences were aligned to multiple genomes using BLAST and regions with conservation in at least the seed region of the most abundant read (nucleotides 2-7) were evaluated for the potential to form hairpins. The similarity between hairpins was evaluated by RNAforester (Lorenz et al., 2011), which takes into account both primary and secondary structure of the sequence, and the highest scoring hairpin above the threshold of 0.2 was considered as a homolog. In addition to

this pre-miRNA level homology search, reads mapping to the hairpin were aligned against all mature sequences from miRBase (Griffiths-Jones et al., 2006; Kozomara et al., 2011) that have the same seed region (nucleotides 2-7) and the best-scoring miRBase miRNA with at least 15 identical bases was considered as potential homolog.

The full miR-Intess analysis results are available at http://www.macgenome.org/pub/ML_140120/

6. ISH.

Whole mount *in situ* hybridization (ISH) with both miRNA probes and *synaptotagmin* probe was carried out by following an earlier described protocol (Pfister et al., 2007) and using young adult worms 4-8 weeks old adults. Pictures were made using Nomarski microscope DIC optics and AxioCam HRC (Zeiss) digital camera.

Synaptotagmin-1 probe was synthesized in a following way: cDNA synthesis was performed using the SuperScript™ III First-Strand Synthesis System (Life Technologies, Cat. No. 18080-051) following the manufacturer's protocol with 2-3 ug of total RNA as template per reaction. Provided oligo(dT) and hexamer random primers were used. Probe template was amplified from cDNA by standard PCR with GoTaq Flexi DNA Polymerase (Promega, Cat. No. M8305). Forward primer used was 5'-AGATTTGAGCTCAGCCTCTG-3' and reverse primer sequence contained T7 promoter: 5'- GGATCCTAATACGACTCACTATAGG-AGCAGGAACACCAGCAATAC-3'.

Digoxigenin (DIG) labelled RNA probe (747 bp in length) was generated using the DIG RNA labeling Mix (Roche, Cat. No. 11277073910) and T7 RNA polymerase (Promega, Cat. No. P2075) following manufacturer's protocol for *in vitro* transcription. Concentration of the probe was measured with Qubit RNA BR assay kit (Invitrogen, Cat. No. Q10211), then probe was diluted in Hybridization Mix (Pfister et al., 2007) to 20 ng/μl, stored at -80°C and used within 4 months. Final concentration of the probe and optimal temperature for hybridization were 1ng/μl and 55°C.

Double-DIG labeled LNA (locked nucleic acid) miRCURY probes for miRNAs were ordered from Exiqon. MiR-7 probe sequence was 5'-UGGAAGACUAGUGAUUUGUUGUC-3' and miR-184 probe sequence was 5'-UGGACGGAGGACUGUUAAGGGC-3'. Both probes were DIG-labeled on both ends. Probes were used in concentration of 2 pM, and miR-184 was also used in concentration of 8 pM (increased probe concentration conditions). Signal development time for the probes was much longer than for normal ISH probes – both probes were developed for 4-5 hours at room temperature. In the case of extended development time reaction with miR-184 probe was developed at room temperature overnight (12 hours).

REFERENCES

Aalto, A.P., and Pasquinelli, A.E. (2012). Small non-coding RNAs mount a silent revolution in gene expression. Curr Opin Cell Biol *24*, 333-340.

Aboobaker, A.A. (2011). Planarian stem cells: a simple paradigm for regeneration. Trends Cell Biol *21*, 304-311.

Aboobaker, A.A., Tomancak, P., Patel, N., Rubin, G.M., and Lai, E.C. (2005). Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. Proc Natl Acad Sci U S A *102*, 18017-022.

Agata, K., Nakajima, E., Funayama, N., Shibata, N., Saito, Y., and Umesono, Y. (2006). Two different evolutionary origins of stem cell systems and their molecular basis. Semin Cell Dev Biol *17*, 503-09.

Almuedo-Castillo, M., Sureda-Gómez, M., and Adell, T. (2012). Wnt signaling in planarians: new answers to old questions. Int J Dev Biol *56*, 53-65.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402

Ambros, V. (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell *113*, 673-76.

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-55.

Andersen, R.A., Berges, R.A., Harrison, P.J., and Watanabe, M.M. (2005). Recipes for freshwater and seawater media; enriched natural seawater media. In: Anderson, R.A. (Ed.), Algal culturing techniques, Elsevier academic press, Burlington, 507.

Axtell, M.J., Westholm, J.O., and Lai, E.C. (2011). Vive la différence: biogenesis and evolution of microRNAs in plants and animals. Genome Biol *12*, 221.

Baguñà, J. (2012). The planarian neoblast: the rambling history of its origin and some current black boxes. Int J Dev Biol *56*, 19-37.

Barroso-delJesus, A., Romero-López, C., Lucena-Aguilar, G., Melen, G.J., Sanchez, L., Ligero, G., Berzal-Herranz, A., and Menendez, P. (2008). Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. Mol Cell Biol *28*, 6609-619.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Berardi, E., Pues, M., Thorrez, L., and Sampaolesi, M. (2012). miRNAs in ESC differentiation. American Journal of Physiology-Heart and Circulatory Physiology *303*, H931-39.

Berezikov, E., Thuemmler, F., van Laake, L.W., Kondova, I., Bontrop, R., Cuppen, E., and Plasterk, R.H. (2006). Diversity of microRNAs in human and chimpanzee brain. Nat Genet *38*, 1375-77.

Berezikov, E., Chung, W.-J., Willis, J., Cuppen, E., and Lai, E.C. (2007). Mammalian mirtron genes. Molecular cell *28*, 328-336.

Berezikov E., Liu N., Flynt A. S., Hodges E., Rooks M., et al. (2010). Nat Genet. In Evolutionary flux of canonical microRNAs and mirtrons in Drosophila.

Berezikov, E. (2011). Evolution of microRNA diversity and regulation in animals. Nat Rev Genet *12*, 846-860.

Berezikov, E., Robine, N., Samsonova, A., Westholm, J.O., Naqvi, A., Hung, J.H., Okamura, K., Dai, Q., Bortolamiol-Becet, D., et al. (2011). Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. Genome Res *21*, 203-215.

Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell *125*, 1111-124.

Bonnet, E., Wuyts, J., Rouzé, P., and Van de Peer, Y. (2004). Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics *20*, 2911-17.

Cai, B., Li, J., Wang, J., Luo, X., Ai, J., Liu, Y., Wang, N., Liang, H., Zhang, M., et al. (2012). microRNA-124 regulates cardiomyocyte differentiation of bone marrow-derived mesenchymal stem cells via targeting STAT3 signaling. Stem Cells *30*, 1746-755.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA *10*, 1957-966.

Candiani, S., Moronti, L., De Pietri Tonelli, D., Garbarino, G., and Pestarino, M. (2011). A study of neural-related microRNAs in the developing amphioxus. Evodevo *2*, 15.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chen, J.F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., and Wang, D.Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet *38*, 228-233.

Choi, E., Choi, E., and Hwang, K.-C. (2013). MicroRNAs as novel regulators of stem cell fate. World J Stem Cells *5*, 172.

Choudhury, N.R., de Lima Alves, F., de Andrés-Aguayo, L., Graf, T., Cáceres, J.F., Rappsilber, J., and Michlewski, G. (2013). Tissue-specific control of brain-enriched miR-7 biogenesis. Genes Dev *27*, 24-38.

Cordes, K.R., Sheehy, N.T., White, M.P., Berry, E.C., Morton, S.U., Muth, A.N., Lee, T.H., Miano, J.M., Ivey, K.N., and Srivastava, D. (2009). miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. Nature *460*, 705-710.

Demircan, T., and Berezikov, E. (2013). The Hippo pathway regulates stem cells during homeostasis and regeneration of the flatworm Macrostomum lignano. Stem Cells Dev *22*, 2174-185.

Egger, B., Ladurner, P., Nimeth, K., Gschwentner, R., and Rieger, R. (2006). The regeneration capacity of the flatworm Macrostomum lignano--on repeated regeneration, rejuvenation, and the minimal size needed for regeneration. Dev Genes Evol *216*, 565-577.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional

regulation by microRNAs: are the answers in sight? Nat Rev Genet *9*, 102-114.

Förstemann, K., Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007). Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. Cell *130*, 287-297.

Friedländer, M.R., Adamidi, C., Han, T., Lebedeva, S., Isenbarger, T.A., Hirst, M., Marra, M., Nusbaum, C., Lee, W.L., et al. (2009). High-resolution profiling and discovery of planarian small RNAs. Proc Natl Acad Sci U S A *106*, 11546-551.

Gentile, L., Cebrià, F., and Bartscherer, K. (2011). The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration. Dis Model Mech *4*, 12-19.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet *10*, 94-108.

Gommans, W.M., and Berezikov, E. (2012). Sample preparation for small RNA massive parallel sequencing. Methods Mol Biol *786*, 167-178

González-Estévez, C., Arseni, V., Thambyrajah, R.S., Felix, D.A., and Aboobaker, A.A. (2009). Diverse miRNA spatial expression patterns suggest important roles in homeostasis and regeneration in planarians. International journal of developmental biology *53*, 493-505.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res *34*, D140-44.

Griffiths-Jones, S., Hui, J.H., Marco, A., and Ronshaugen, M. (2011). MicroRNA evolution by arm switching. EMBO Rep *12*, 172-77.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B.J., Chiang, H.R., King, N., Degnan, B.M., Rokhsar, D.S., and Bartel, D.P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature *455*, 1193-97.

Guo, L., Zhao, R.C., and Wu, Y. (2011). The role of microRNAs in self-renewal and differentiation of mesenchymal stem cells. Exp Hematol *39*, 608-616.

Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008). Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science *319*, 323-27.

Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Dev Growth Differ *48*, 371-380.

Heinrich, E.M., and Dimmeler, S. (2012). MicroRNAs and stem cells: control of pluripotency, reprogramming, and lineage commitment. Circ Res *110*, 1014-022.

Höchsmann, M., Voss, B., and Giegerich, R. (2004). Pure multiple RNA secondary structure alignments: a progressive profile approach. IEEE/ACM Trans Comput Biol Bioinform *1*, 53-62

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. Dev Cell *5*, 351-58.

Howard, L., Kane, N.M., Milligan, G., and Baker, A.H. (2011). MicroRNAs regulating cell pluripotency and vascular differentiation. Vascul Pharmacol *55*, 69-78.

Huang, J., Hao, P., Chen, H., Hu, W., Yan, Q., Liu, F., and Han, Z.G. (2009). Genome-wide identification of Schistosoma japonicum microRNAs using a deep-sequencing approach. PLoS One *4*, e8206.

Ivey, K.N., Muth, A., Arnold, J., King, F.W., Yeh, R.F., Fish, J.E., Hsiao, E.C., Schwartz, R.J., Conklin, B.R., et al. (2008). MicroRNA regulation of cell lineages in mouse and human embryonic stem cells. Cell Stem Cell *2*, 219-229.

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D.B., and Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution *67*, 3233-242.

Jiao, A.L., and Slack, F.J. (2013). RNA-mediated gene activation. Epigenetics *9*

Kane, N.M., Meloni, M., Spencer, H.L., Craig, M.A., Strehl, R., Milligan, G., Houslay, M.D., Mountford, J.C., Emanueli, C., and Baker, A.H. (2010). Derivation of endothelial cells from human embryonic stem cells by directed differentiation: analysis of microRNA and angiogenesis in vitro and in vivo. Arterioscler Thromb Vasc Biol *30*, 1389-397.

Kapinas, K., Kessler, C.B., and Delany, A.M. (2009). miR-29 suppression of osteonectin in osteoblasts: regulation during differentiation and by canonical Wnt signaling. J Cell Biochem *108*, 216-224.

Karres, J.S., Hilgers, V., Carrera, I., Treisman, J., and Cohen, S.M. (2007). The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in Drosophila. Cell *131*, 136-145.

Kawahara, H., Imai, T., and Okano, H. (2012). MicroRNAs in Neural Stem Cells and Neurogenesis. Front Neurosci *6*, 30.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res *39*, D152-57.

Kwon, C., Han, Z., Olson, E.N., and Srivastava, D. (2005). MicroRNA1 influences cardiac differentiation in Drosophila and regulates Notch signaling. Proc Natl Acad Sci U S A *102*, 18986-991.

Ladurner, P., Pfister, D., Seifarth, C., Schärer, L., Mahlknecht, M., Salvenmoser, W., Gerth, R., Marx, F., and Rieger, R. (2005). Production and characterisation of cell- and tissue-specific monoclonal antibodies for

the flatworm Macrostomum sp. Histochem Cell Biol *123*, 89-104.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Lang, M.F., and Shi, Y. (2012). Dynamic Roles of microRNAs in Neurogenesis. Front Neurosci *6*, 71.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J *23*, 4051-060.

Li, Y.-Q., Zeng, A., Han, X.-S., Wang, C., Li, G., Zhang, Z.-C., Wang, J.-Y., Qin, Y.-W., and Jing, Q. (2011). Argonaute-2 regulates the proliferation of adult stem cells in planarian. Cell Res *21*, 1750-54.

Li, Z., Hassan, M.Q., Jafferji, M., Aqeilan, R.I., Garzon, R., Croce, C.M., van Wijnen, A.J., Stein, J.L., Stein, G.S., and Lian, J.B. (2009). Biological functions of miR-29b contribute to positive regulation of osteoblast differentiation. J Biol Chem *284*, 15676-684.

Lian, J.B., Stein, G.S., van Wijnen, A.J., Stein, J.L., Hassan, M.Q., Gaur, T., and Zhang, Y. (2012). MicroRNA control of bone formation and homeostasis. Nat Rev Endocrinol *8*, 212-227.

Lin, C.H., Jackson, A.L., Guo, J., Linsley, P.S., and Eisenman, R.N. (2009). Myc-regulated microRNAs attenuate embryonic stem cell differentiation. EMBO J *28*, 3157-170.

Liu, C., Teng, Z.Q., Santistevan, N.J., Szulwach, K.E., Guo, W., Jin, P., and Zhao, X. (2010). Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation. Cell Stem Cell *6*, 433-444.

Liu, N., Okamura, K., Tyler, D.M., Phillips, M.D., Chung, W.J., and Lai, E.C. (2008). The evolution and functional diversification of animal microRNA genes. Cell Res *18*, 985-996.

Liu, P.Q., Chan, E.M., Cost, G.J., Zhang, L., Wang, J., Miller, J.C., Guschin, D.Y., Reik, A., Holmes, M.C., et al. (2010). Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. Biotechnol Bioeng *106*, 97-105.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol Biol *6*, 26.

Lu, J., Shen, Y., Wu, Q., Kumar, S., He, B., Shi, S., Carthew, R.W., Wang, S.M., and Wu, C.I. (2008). The birth and death of microRNA genes in Drosophila. Nat Genet *40*, 351-55.

Lu, Y.C., Smielewska, M., Palakodeti, D., Lovci, M.T., Aigner, S., Yeo, G.W., and Graveley, B.R. (2009). Deep sequencing identifies new and regulated microRNAs in Schmidtea mediterranea. RNA *15*, 1483-491.

Lu, Y.C., Smielewska, M., Palakodeti, D., Lovci, M.T., Aigner, S., Yeo, G.W., and Graveley, B.R. (2009). Deep sequencing identifies new and regulated microRNAs in Schmidtea mediterranea. RNA *15*, 1483-491.

Macarthur, B.D., Ma'ayan, A., and Lemischka, I.R. (2009). Systems biology of stem cell fate and cellular reprogramming. Nat Rev Mol Cell Biol *10*, 672-681.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625-635.

Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., and Proudfoot, N.J. (2008). Primary microRNA transcripts are processed co-transcriptionally. Nat Struct Mol Biol *15*, 902-09.

Mouton, S., Willems, M., Braeckman, B.P., Egger, B., Ladurner, P., Schärer, L., and Borgonie, G. (2009). The free-living flatworm Macrostomum lignano: a new model organism for ageing research. Exp Gerontol *44*, 243-49.

De Mulder, K., Kuales, G., Pfister, D., Egger, B., Seppi, T., Eichberger, P., Borgonie, G., and Ladurner, P. (2010). Potential of Macrostomum lignano to recover from gamma-ray irradiation. Cell Tissue Res *339*, 527-542.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., et al. (2009). Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. Dev Biol *334*, 198-212.

Mullen, A.C., Orlando, D.A., Newman, J.J., Lovén, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-β signaling. Cell *147*, 565-576.

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci U S A *102*, 12135-140.

Newmark, P.A., and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev Biol *220*, 142-153.

Newmark, P.A., Reddien, P.W., Cebrià, F., and Sánchez Alvarado, A. (2003). Ingestion of bacterially expressed

double-stranded RNA inhibits gene expression in planarians. Proc Natl Acad Sci U S A *100 Suppl 1*, 11861-65.

Nimeth, K.T., Mahlknecht, M., Mezzanato, A., Peter, R., Rieger, R., and Ladurner, P. (2004). Stem cell dynamics during growth, feeding, and starvation in the basal flatworm Macrostomum sp. (Platyhelminthes). Dev Dyn *230*, 91-99.

O'Connell, R.M., Taganov, K.D., Boldin, M.P., Cheng, G., and Baltimore, D. (2007). MicroRNA-155 is induced during the macrophage inflammatory response. Proc Natl Acad Sci U S A *104*, 1604-09.

Ohtani, K., and Dimmeler, S. (2011). Control of cardiovascular differentiation by microRNAs. Basic Res Cardiol *106*, 5-11.

Palakodeti, D., Smielewska, M., and Graveley, B.R. (2006). MicroRNAs from the Planarian Schmidtea mediterranea: a model system for stem cell biology. RNA *12*, 1640-49.

Pan, G., and Thomson, J.A. (2007). Nanog and transcriptional networks in embryonic stem cell pluripotency. Cell Res *17*, 42-49.

Papagiannakopoulos, T., and Kosik, K.S. (2009). MicroRNA-124: micromanager of neurogenesis. Cell Stem Cell *4*, 375-76.

Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., and Sánchez Alvarado, A. (2009). Formaldehyde-based whole-mount in situ hybridization method for planarians. Dev Dyn *238*, 443-450.

Petersen, C.P., and Reddien, P.W. (2008). Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. Science *319*, 327-330.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. (2007). The exceptional stem cell system of Macrostomum lignano: screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. Front Zool *4*, 9.

Qin, Y.F., Zhao, J.M., Bao, Z.X., Zhu, Z.Y., Mai, J., Huang, Y.B., Li, J.B., Chen, G., Lu, P., et al. (2012). Identification of small non-coding RNAs in the planarian Dugesia japonica via deep sequencing. Genomics *99*, 315-321.

Reddien, P.W., and Sánchez Alvarado, A. (2004). Fundamentals of planarian regeneration. Annu Rev Cell Dev Biol *20*, 725-757.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-06.

Resch, A.M., and Palakodeti, D. (2012). Small RNA pathways in Schmidtea mediterranea. International Journal of Developmental Biology *56*, 67.

Resch, A.M., Palakodeti, D., Lu, Y.C., Horowitz, M., and Graveley, B.R. (2012). Transcriptome analysis reveals strain-specific and conserved stemness genes in Schmidtea mediterranea. PLoS One *7*, e34447.

Rieger, R.M., Gehlen, M., Haszprunar, G., Holmlund, M., Legniti, A., Salvenmoser, W., and Tyler, S. (1988). Laboratory cultures of marine Macrostomida (Turbellaria). Fortschr Zool *36*, 523.

Rink, J.C. (2013). Stem cell systems and regeneration in planaria. Dev Genes Evol *223*, 67-84.

Rink, J.C. (2013). Stem cell systems and regeneration in planaria. Dev Genes Evol *223*, 67-84.

Ritchie, W., Legendre, M., and Gautheret, D. (2007). RNA stem--loops: To be or not to be cleaved by RNAse III. RNA *13*, 457-462.

Robb, S.M., Ross, E., and Sánchez Alvarado, A. (2008). SmedGD: the Schmidtea mediterranea genome database. Nucleic Acids Res *36*, D599-D606.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res *14*, 1902-910.

Rouhana, L., Shibata, N., Nishimura, O., and Agata, K. (2010). Different requirements for conserved post-transcriptional regulators in planarian regeneration and stem cell maintenance. Dev Biol *341*, 429-443.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83-86

Ruby, J.G., Stark, A., Johnston, W.K., Kellis, M., Bartel, D.P., and Lai, E.C. (2007). Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res *17*, 1850-864.

Sasidharan, V., Lu, Y.C., Bansal, D., Dasari, P., Poduval, D., Seshasayee, A., Resch, A.M., Graveley, B.R., and Palakodeti, D. (2013). Identification of neoblast- and regeneration-specific miRNAs in the planarian Schmidtea mediterranea. RNA *19*, 1394-1404.

Sánchez Alvarado, A. (2004). Regeneration and the need for simpler model organisms. Philos Trans R Soc Lond B Biol Sci *359*, 759-763.

Sánchez Alvarado, A. (2007). Stem cells and the Planarian Schmidtea mediterranea. Comptes rendus biologies *330*, 498-503.

Sánchez Alvarado, A., and Newmark, P.A. (1999). Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proc Natl Acad Sci U S A *96*, 5049-054.

Schwab, R., and Voinnet, O. (2009). miRNA processing turned upside down. EMBO J *28*, 3633-34.

Schwarz, D.S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.

Silva, J., and Smith, A. (2008). Capturing pluripotency. Cell *132*, 532-36.

Simanov, D., Mellaart-Straver, I., Sormacheva, I., and Berezikov, E. (2012). The Flatworm Macrostomum lignano Is a Powerful Model Organism for Ion Channel and Stem Cell Research. Stem Cells Int *2012*, 167265.

Simões, M.C., Lee, J., Djikeng, A., Cerqueira, G.C., Zerlotini, A., da Silva-Pereira, R.A., Dalby, A.R., LoVerde, P., El-Sayed, N.M., and Oliveira, G. (2011). Identification of Schistosoma mansoni microRNAs. BMC Genomics *12*, 47.

Song, X., Wang, M., Chen, Y.P., Wang, H., Han, P., and Sun, H. (2013). Prediction of pre-miRNA with multiple stem-loops using pruning algorithm. Comput Biol Med *43*, 409-416.

Sugatani, T., and Hruska, K.A. (2007). MicroRNA-223 is a key factor in osteoclast differentiation. J Cell Biochem *101*, 996-99.

Suh, M.R., Lee, Y., Kim, J.Y., Kim, S.K., Moon, S.H., Lee, J.Y., Cha, K.Y., Chung, H.M., Yoon, H.S., et al. (2004). Human embryonic stem cells express a unique set of microRNAs. Dev Biol *270*, 488-498.

Szulwach, K.E., Li, X., Smrt, R.D., Li, Y., Luo, Y., Lin, L., Santistevan, N.J., Li, W., Zhao, X., and Jin, P. (2010). Cross talk between microRNA and epigenetic regulation in adult neurogenesis. J Cell Biol *189*, 127-141.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Thatcher, E.J., and Patton, J.G. (2010). Small RNAs have a big impact on regeneration. RNA Biol *7*, 333-38.

Tomari, Y., Du, T., and Zamore, P.D. (2007). Sorting of Drosophila small silencing RNAs. Cell *130*, 299-308.

Tran, N., and Hutvagner, G. (2013). Biogenesis and the regulation of the maturation of miRNAs. Essays Biochem *54*, 17-28.

Umesono, Y., Watanabe, K., and Agata, K. (1997). A planarian orthopedia homolog is specifically expressed in the branch region of both the mature and regenerating brain. Dev Growth Differ *39*, 723-27.

Verdoodt, F., Bert, W., Couvreur, M., De Mulder, K., and Willems, M. (2012). Proliferative response of the stem cell system during regeneration of the rostrum in Macrostomum lignano (Platyhelminthes). Cell Tissue Res *347*, 397-406.

Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. Science *332*, 811-16.

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Blelloch, R. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet *40*, 1478-483.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet *39*, 380-85.

de Wit, E., Linsen, S.E., Cuppen, E., and Berezikov, E. (2009). Repertoire and evolution of miRNA genes in four divergent nematode species. Genome Res *19*, 2064-074.

Wu, J., Bao, J., Wang, L., Hu, Y., and Xu, C. (2011). MicroRNA-184 downregulates nuclear receptor corepressor 2 in mouse spermatogenesis. BMC Dev Biol *11*, 64.

Xiao, C., Calado, D.P., Galler, G., Thai, T.H., Patterson, H.C., Wang, J., Rajewsky, N., Bender, T.P., and Rajewsky, K. (2007). MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb. Cell *131*, 146-159.

Xiao, J., Liang, D., Zhang, H., Liu, Y., Zhang, D., Liu, Y., Pan, L., Chen, X., Doevendans, P.A., et al. (2012). MicroRNA-204 is required for differentiation of human-derived cardiomyocyte progenitor cells. J Mol Cell Cardiol *53*, 751-59.

Xu, Z., Chen, M., Ren, Z., Zhang, N., Xu, H., Liu, X., Tian, G., Song, L., and Yang, H. (2013). Deep sequencing identifies regulated small RNAs in Dugesia japonica. Mol Biol Rep *40*, 4075-081.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-920.

Zhao, Y., Ransom, J.F., Li, A., Vedantham, V., von Drehle, M., Muth, A.N., Tsuchihashi, T., McManus, M.T., Schwartz, R.J., and Srivastava, D. (2007). Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. Cell *129*, 303-317.

SUPPLEMENTARY

Supplementary Table 1. All miRNAs identified in *M. lignano*.

miR-Intess ID	Best homolog miRBase ID	chr	from	to	strand	MIRCTRL	MIRIRR	MIRCTRL*	MIRIRR*
sblock19814	let-7b	chrS5	12202548	12202647	-1	54673	183493	1916	970
sblock13518	let-7d	chrD4	941559	941711	1	55449	75751	27	13
sblock2594	let-7e	chrD10	8981932	8982081	1	9916	8948	27	13
sblock19192	mir-1	chrS3	168247	168336	1	3185	1020	5	0
sblock13077	miR-1c-3p	chrD3	1070766	1071032	-1	3185	1020	2	0
sblock19451	miR-1c-3p	chrS3	21901622	21901755	1	5369	732	4	0
sblock2141	miR-1c-3p	chrD10	2121482	2121773	-1	3185	1020	2	0
sblock18340	mir-2b	chrS1	8883965	8884044	1	803	97	25	256
sblock9419	mir-2b	chrD16	11384903	11384981	1	803	97	25	256
sblock18341	mir-2a-1	chrS1	8884043	8884177	1	922	32	0	4
sblock19743	mir-2a-1	chrS5	2701744	2701825	-1	922	32	0	4
sblock9420	mir-2a-1	chrD16	11384984	11385120	1	922	32	0	4
sblock19666	miR-7-5p	chrS4	12436872	12437032	1	9887	346228	2	0
sblock11336	mir-8a	chrD17	16156842	16156949	1	655	28	13	1
sblock11337	mir-8a	chrD17	16157294	16157393	1	655	28	13	1
sblock6458	mir-8a	chrD14	13707329	13707448	-1	756	33	1930	70
sblock18671	mir-9	chrS2	20825682	20825813	-1	34	6	7	4
sblock7686	miR-10b	chrD15	18091206	18091355	-1	33	56	13	1
sblock9297	miR-10b	chrD16	22232253	22232406	-1	33	56	13	1
sblock9298	miR-10b	chrD16	22232797	22232946	-1	33	56	13	1
sblock2598	miR-10-3p	chrD10	9903745	9903849	1	13	2	1	2
sblock2599	miR-10-3p	chrD10	9903972	9904097	1	13	2	1	2
sblock13364	miR-10-3p	chrD4	2350062	2350187	-1	13	2	1	2
sblock19541	miR-22a-3p	chrS4	13414521	13414642	-1	12491	9557	1530	1204
sblock13339	miR-29c-3p	chrD3	22859294	22859478	1	322	291	6	0
sblock19947	miR-29d-3p	chrS5	9022163	9022272	1	39	20	2	0
sblock17917	miR-33b	chrD9	18872951	18873150	1	1287	124	4	0
sblock18471	miR-44-3p	chrS1	24996639	24996756	1	20461	35051	74	80
sblock18335	mir-46	chrS1	8305417	8305507	1	356	120	41	22
sblock20606	miR-49-3p	chrS6	15164714	15164820	1	2263	122	336	105
sblock18922	miR-50	chrS2	23673446	23673581	1	43	0	3	0
sblock18339	mir-71b	chrS1	8883853	8883940	1	41031	10462	556	203
sblock6710	mir-71b	chrD14	8598646	8598750	1	41031	10462	87	6
sblock18518	mir-72	chrS2	5257304	5257497	-1	3462	1112	8	0
sblock13734	miR-72-5p	chrD5	11896340	11896463	-1	3436	1087	9	0
sblock20587	miR-87-3p	chrS6	13586357	13586505	1	168	13	8	2
sblock12166	miR-87-3p	chrD18	13965435	13965550	-1	137	13	59	45
sblock13249	miR-87-3p	chrD3	8028511	8028626	1	137	13	59	45
sblock13332	miR-87-3p	chrD3	21487794	21487909	1	137	13	59	45

Supplementary Table 1. All miRNAs identified in *M. lignano*. (*Continued*)

miR-Intess ID	Best homolog miRBase ID	chr	from	to	strand	MIRCTRL	MIRIRR	MIRCTRL*	MIRIRR*
sblock2151	miR-87-3p	chrD10	3015656	3015790	-1	137	13	59	45
sblock20438	miR-92-5p	chrS6	25697611	25697725	-1	145	1339	2	2
sblock12796	miR-92-5p	chrD2	19257801	19257886	-1	800	1362	2	2
sblock18402	miR-92-5p	chrS1	15094129	15094210	1	800	1362	2	2
sblock19316	miR-92c-3p	chrS3	12081048	12081192	1	3419	296	20	39
sblock19546	miR-96a	chrS4	13936349	13936477	-1	58363	5962	18	16
sblock19584	miR-124-3p	chrS4	23888387	23888636	-1	81	40	4	0
sblock19918	ola-miR-125c	chrS5	4100055	4100178	1	49403	42329	393	56
sblock13128	mir-133a	chrD3	10455745	10455828	-1	102	17	1	0
sblock13038	miR-182b-5p	chrD2	20565349	20565464	1	29542	11562	1	0
sblock13334	miR-182b-5p	chrD3	22003663	22003806	1	29544	11568	14	129
sblock20409	miR-182b-5p	chrS6	22704113	22704220	-1	29543	11562	116	39
sblock3707	miR-182b-5p	chrD11	11429140	11429253	1	29542	11562	1	0
sblock19138	miR-184-3p	chrS3	18131306	18131455	-1	954	309	53	31
sblock13174	mir-184a	chrD3	18445299	18445414	-1	61511	12884	16	8
sblock20563	miR-184	chrS6	11264915	11265057	1	1590	253	164	129
sblock5036	miR-193a-5p	chrD12	22691949	22692048	-1	52	17	10	10
sblock5037	miR-193a-5p	chrD12	22692343	22692442	-1	52	17	10	10
sblock5035	miR-193a-5p	chrD12	22691545	22691615	-1	52	17	10	10
sblock2000	mir-219	chrD1	19198156	19198271	1	175	313	7	8
sblock6327	mir-219	chrD13	14133048	14133162	1	175	313	7	8
sblock12783	mir-236	chrD2	17300261	17300360	-1	655	28	13	1
sblock17468	mir-236	chrD8	22920127	22920226	1	755	32	45	26
sblock2920	mir-236	chrD10	19545602	19545701	1	755	32	45	26
sblock12650	mir-252a	chrD2	4398077	4398177	-1	669	237	10	0
sblock17657	mir-252a	chrD9	10081312	10081413	-1	669	237	10	0
sblock17658	mir-252a	chrD9	10081677	10081781	-1	669	237	10	0
sblock13164	mir-277	chrD3	16305411	16305535	-1	1637	195	205	161
sblock17744	mir-277	chrD9	23481854	23481975	-1	1639	195	204	161
sblock19567	miR-277	chrS4	19907628	19907744	-1	14064	9023	2357	1623
sblock19380	miR-277c-3p	chrS3	16108109	16108247	1	4336	450	47	27
sblock18780	miR-287	chrS2	7534634	7534752	1	787	517	75	12
sblock18464	miR-375-3p	chrS1	24198604	24198730	1	23403	12032	153	18
sblock19255	miR-745b	chrS3	6208304	6208430	1	63	2028	31	17
sblock19543	miR-745	chrS4	13414829	13414947	-1	11670	8309	567	21
sblock19966	miR-980-3p	chrS5	9941287	9941436	1	7970	3208	48	25
sblock19387	miR-981	chrS3	16601647	16601789	1	1150	331	3	0
sblock2128	miR-981	chrD10	299634	299724	-1	1765	828	2	8
sblock2131	miR-981	chrD10	300084	300174	-1	1765	828	2	8
sblock20164	miR-993b-3p	chrS6	2240186	2240358	-1	2771	706	9	28
sblock6462	miR-993	chrD14	14150641	14150754	-1	492	31	13	13

Supplementary Table 1. All miRNAs identified in *M. lignano*. (*Continued*)

miR-Intess ID	Best homolog miRBase ID	chr	from	to	strand	MIRCTRL	MIRIRR	MIRCTRL*	MIRIRR*
sblock18500	miR-996	chrS2	2366244	2366358	-1	7699	1689	60	0
sblock10589	miR-996	chrD17	19072526	19072606	-1	1274	2344	13	5
sblock10590	miR-996	chrD17	19073070	19073150	-1	1274	2344	13	5
sblock11994	miR-996	chrD18	4734003	4734085	-1	1274	2344	13	5
sblock11995	miR-996	chrD18	4734375	4734457	-1	1274	2344	13	5
sblock12191	miR-996	chrD18	15295673	15295768	-1	1274	2344	13	5
sblock18891	miR-996	chrS2	21620641	21620721	1	1274	2344	13	5
sblock5033	miR-996	chrD12	22549313	22549395	-1	1274	2344	13	5
sblock6573	miR-996	chrD14	18716390	18716471	-1	1274	2344	13	5
sblock778	miR-996	chrD1	13261715	13261810	-1	1274	2344	13	5
sblock8706	miR-996	chrD16	10000600	10000680	-1	1274	2344	13	5
sblock10345	miR-1175-3p	chrD17	12560070	12560194	-1	595	9	105	5
sblock10346	miR-1175-3p	chrD17	12560494	12560630	-1	595	9	105	5
sblock19133	miR-1175-3p	chrS3	17735520	17735668	-1	108	5	203	5
sblock19439	miR-1175-3p	chrS3	21280555	21280706	1	595	9	103	3
sblock20552	miR-1175-3p	chrS6	9830505	9830650	1	595	9	108	5
sblock2516	miR-1175-3p	chrD10	22637253	22637392	-1	595	9	105	5
sblock3468	miR-1175-3p	chrD11	23214703	23214852	-1	108	5	203	5
sblock7725	miR-1175-3p	chrD15	20747686	20747851	-1	595	9	105	5
sblock7726	miR-1175-3p	chrD15	20748217	20748366	-1	595	9	105	5
sblock18396	miR-1989	chrS1	14645530	14645685	1	1665	240	32	40
sblock19147	miR-1992	chrS3	19454030	19454161	-1	8	1	4	0
sblock19030	miR-1992	chrS3	6946203	6946293	-1	9	5	2	2
sblock13027	miR-1992	chrD2	18946376	18946486	1	9	5	2	2
sblock20293	miR-1993	chrS6	10893641	10893783	-1	462	28	0	4
sblock19461	mir-2001	chrS3	23449563	23449645	1	150	29	2	0
sblock19113	miR-2162-3p	chrS3	15369880	15369984	-1	127	22	3	16
sblock19586	miR-3177-5p	chrS4	24398537	24398614	-1	14169	226	118	18
sblock17737	miR-4503	chrD9	22433313	22433398	-1	4	0	1	1
sblock17738	miR-4503	chrD9	22433770	22433857	-1	4	0	1	1
sblock3695	miR-4503	chrD11	9295600	9295687	1	4	0	1	1
sblock19951	miR-5342-3p	chrS5	9348586	9348738	1	582351	219255	14	5
sblock12032	miR-5453	chrD18	6993842	6993922	-1	0	0	0	1
sblock19237	miR-5706	chrS3	5027311	5027509	1	15	26	6	2
sblock6340	miR-5706	chrD13	16544531	16544730	1	15	26	6	2
sblock19353	miR-6673-3p	chrS3	14600594	14600677	1	581	300	12	24
sblock18398	miR-6958-3p	chrS1	14645716	14645859	1	335	134	47	38
sblock2954	miR-6958-3p	chrD10	21566759	21566902	1	335	134	47	38
sblock13376	novel	chrD4	4606990	4607089	-1	2816	2920	39	61
sblock18559	novel	chrS2	9135155	9135273	-1	326	177	4	8
sblock18771	novel	chrS2	6362139	6362301	1	5047	85	96	49

Supplementary Table 1. All miRNAs identified in *M. lignano*. (*Continued*)

miR-Intess ID	Best homolog miRBase ID	chr	from	to	strand	MIRCTRL	MIRIRR	MIRCTRL*	MIRIRR*
sblock19434	novel	chrS3	20805659	20805831	1	40	85	143	55
sblock19587	novel	chrS4	24398671	24398773	-1	1084	1372	17	0
sblock19588	novel	chrS4	24398800	24398899	-1	1155	28	134	24
sblock19777	novel	chrS5	6387619	6387748	-1	18	1	1	0
sblock19852	novel	chrS5	20473665	20473781	-1	201	173	12	0
sblock20297	novel	chrS6	11261467	11261606	-1	31202	20727	83	44
sblock20456	novel	chrS6	28720476	28720632	-1	90	44	34	5
sblock6900	novel	chrD14	23781830	23781951	1	1	3	1	0
sblock11160	candidate	chrD17	8346205	8346272	1	0	1	0	0
sblock12097	candidate	chrD18	8631154	8631231	-1	0	0	0	1
sblock13018	candidate	chrD2	18030484	18030574	1	0	0	0	1
sblock13062	candidate	chrD2	23549498	23549582	1	5	0	1	0
sblock13074	candidate	chrD3	1044285	1044396	-1	1	0	0	0
sblock13265	candidate	chrD3	10514709	10514795	1	1	0	0	0
sblock13271	candidate	chrD3	10992435	10992584	1	0	1	0	0
sblock13626	candidate	chrD4	19753898	19754056	1	1	0	1	0
sblock13674	candidate	chrD5	4567179	4567252	-1	1	4	0	0
sblock1371	candidate	chrD1	3079510	3079596	1	1	0	0	0
sblock13924	candidate	chrD6	3058630	3058787	-1	0	0	0	1
sblock14718	candidate	chrD7	18900491	18900559	1	1	4	0	0
sblock15038	candidate	chrD8	237108	237227	-1	4	0	1	2
sblock17459	candidate	chrD8	22468823	22468898	1	0	0	1	0
sblock17603	candidate	chrD9	5406671	5406763	-1	0	0	2	1
sblock17619	candidate	chrD9	6769245	6769361	-1	1	0	1	0
sblock17867	candidate	chrD9	12325892	12325970	1	0	0	1	2
sblock18183	candidate	chrS1	22012299	22012366	-1	0	1	0	0
sblock18369	candidate	chrS1	11744462	11744568	1	0	0	0	0
sblock18536	candidate	chrS2	7037847	7037946	-1	4	0	0	0
sblock18748	candidate	chrS2	2586854	2586952	1	3	5	0	0
sblock18789	candidate	chrS2	8912371	8912545	1	1	0	0	0
sblock18818	candidate	chrS2	11032163	11032317	1	1	0	0	0
sblock19472	candidate	chrS4	139175	139299	-1	179	79	24	20
sblock19510	candidate	chrS4	6284833	6284929	-1	4	0	0	0
sblock19511	candidate	chrS4	6285603	6285689	-1	5	0	1	0
sblock19736	candidate	chrS5	2419758	2419842	-1	0	0	0	0
sblock20115	candidate	chrS6	211694	211831	-1	3	3	2	1
sblock20208	candidate	chrS6	4026364	4026442	-1	0	0	2	1
sblock20263	candidate	chrS6	8257969	8258047	-1	0	0	0	1
sblock20656	candidate	chrS6	18552857	18553106	1	2	3	0	3
sblock2075	candidate	chrD1	22478360	22478446	1	0	0	0	1
sblock2162	candidate	chrD10	5588393	5588542	-1	0	1	0	0

Supplementary Table 1. All miRNAs identified in *M. lignano*. (*Continued*)

miR-Intess ID	Best homolog miRBase ID	chr	from	to	strand	MIRCTRL	MIRIRR	MIRCTRL*	MIRIRR*
sblock2949	candidate	chrD10	21152316	21152417	1	4	4	8	7
sblock2981	candidate	chrD10	23722325	23722411	1	0	0	2	1
sblock4811	candidate	chrD12	11853859	11853958	-1	1	1	0	0
sblock5619	candidate	chrD12	10163208	10163290	1	2	1	0	0
sblock5620	candidate	chrD12	10163747	10163829	1	2	1	0	0
sblock6037	candidate	chrD13	7923556	7923698	-1	0	0	0	0
sblock6039	candidate	chrD13	7924375	7924508	-1	0	0	0	0
sblock6425	candidate	chrD14	4502704	4502814	-1	1	4	1	0
sblock6426	candidate	chrD14	4503074	4503168	-1	1	4	1	0
sblock6427	candidate	chrD14	5130028	5130179	-1	961	150	7	9
sblock6466	candidate	chrD14	14890550	14890699	-1	0	1	0	0
sblock755	candidate	chrD1	12449894	12449972	-1	0	0	1	0
sblock7696	candidate	chrD15	18350475	18350561	-1	0	0	2	1
sblock8043	candidate	chrD15	5823734	5823838	1	0	1	0	0
sblock8204	candidate	chrD15	7796581	7796648	1	0	1	0	0
sblock8205	candidate	chrD15	7797011	7797078	1	0	1	0	0
sblock8624	candidate	chrD16	4651938	4652089	-1	961	150	7	9
sblock8648	candidate	chrD16	9058072	9058136	-1	1	0	0	0
sblock868	candidate	chrD1	16660106	16660202	-1	0	0	0	1
sblock9374	candidate	chrD16	9273326	9273403	1	0	1	1	0
sblock9608	candidate	chrD16	15251675	15251780	1	0	1	0	0
sblock9888	candidate	chrD16	22583896	22584008	1	7	43	2	1
sblock9940	candidate	chrD17	4216827	4216989	-1	0	1	0	0

MiRBase – related, novel and candidate miRNAs are present in the table (see "Best homolog miRBase" column). "chr" – is chromosome number, and "from" and "to" define position on chromosome, given miRNA maps to. "MIRCTRL" and "MIRIRR" – total number of reads for given miRNA in control and irradiated libraries respectively. "MIRCTRL*" and "MIRIRR*" – number of reads for star miRNAs in control and irradiated libraries.

Supplementary Table 2. Clustered miRNAs.

cluster	ID	Location	Distance
cluster 1: 2 miRNAs	cte-miR-981	chrD10:299634:299724:-1	
	cte-miR-981	chrD10:300084:300174:-1	359
cluster 2: 2 miRNAs	sja-miR-10-3p	chrD10:9903745:9903849:1	
	sja-miR-10-3p	chrD10:9903972:9904097:1	122
cluster 3: 3 miRNAs	mmu-miR-193a-5p	chrD12:22691545:22691615:-1	
	mmu-miR-193a-5p	chrD12:22691949:22692048:-1	333
	mmu-miR-193a-5p	chrD12:22692343:22692442:-1	294
cluster 5: 2 miRNAs	prd-miR-7899-5p	chrD13:7923556:7923698:-1	
	prd-miR-7899-5p	chrD13:7924375:7924508:-1	676
cluster 7: 2 miRNAs	lgi-miR-1175-3p	chrD15:20747686:20747851:-1	
	lgi-miR-1175-3p	chrD15:20748217:20748366:-1	365
cluster 8: 2 miRNAs	mghv-miR-M1-3-3p	chrD15:7796581:7796648:1	
	mghv-miR-M1-3-3p	chrD15:7797011:7797078:1	362
cluster 9: 2 miRNAs	cte-miR-10b	chrD16:22232253:22232406:-1	
	cte-miR-10b	chrD16:22232797:22232946:-1	390
cluster 10: 2 miRNAs	sme-mir-2b	chrD16:11384903:11384981:1	
	sme-mir-2a-1	chrD16:11384984:11385120:1	2
cluster 11: 2 miRNAs	lgi-miR-1175-3p	chrD17:12560070:12560194:-1	
	lgi-miR-1175-3p	chrD17:12560494:12560630:-1	299
cluster 12: 2 miRNAs	api-miR-996	chrD17:19072526:19072606:-1	
	api-miR-996	chrD17:19073070:19073150:-1	463
cluster 13: 2 miRNAs	sme-mir-8a	chrD17:16156842:16156949:1	
	sme-mir-8a	chrD17:16157294:16157393:1	344
cluster 14: 2 miRNAs	api-miR-996	chrD18:4734003:4734085:-1	
	api-miR-996	chrD18:4734375:4734457:-1	289
cluster 15: 2 miRNAs	cte-mir-252a	chrD9:10081312:10081413:-1	
	cte-mir-252a	chrD9:10081677:10081781:-1	263
cluster 16: 2 miRNAs	hsa-miR-4503	chrD9:22433313:22433398:-1	
	hsa-miR-4503	chrD9:22433770:22433857:-1	371
cluster 17: 3 miRNAs	sma-mir-71b	chrS1:8883853:8883940:1	
	sme-mir-2b	chrS1:8883965:8884044:1	24
	sme-mir-2a-1	chrS1:8884043:8884177:1	-2
cluster 18: 2 miRNAs	lgi-miR-1989	chrS1:14645530:14645685:1	
	mmu-miR-6958-3p	chrS1:14645716:14645859:1	30
cluster 19: 2 miRNAs	dps-miR-2507a	chrS4:6284833:6284929:-1	
	sme-miR-2160-5p	chrS4:6285603:6285689:-1	673
cluster 20: 2 miRNAs	pma-miR-22a-3p	chrS4:13414521:13414642:-1	
	egr-miR-745	chrS4:13414829:13414947:-1	186
cluster 21: 3 miRNAs	hsa-miR-3177-5p	chrS4:24398537:24398614:-1	
	sblock19587_novel	chrS4:24398671:24398773:-1	56
	sblock19588_novel	chrS4:24398800:24398899:-1	26

miRNA clusters in *M. lignano*. miRNAs that lay within 500 bases in the genome are considered clustered. The distance is calculated as the interval between the end of one hairpin and the beginning of the next hairpin.

CHAPTER 7

GENERAL DISCUSSION

RISE OF THE NEW MODEL ORGANISM

The last two decades were marked by great discoveries and significant breakthroughs in different biological disciplines, and stem cell biology is a great example of the field developing in an explosive manner. Discovery of stem cell niches (Xie and Spradling, 2000; Kiger et al., 2000), general improvement of our knowledge of the molecular mechanisms underlying regeneration (Brockes and Kumar, 2008; King and Newmark, 2012; Levin, 2012) and successful reprogramming of somatic cells into pluripotent stem state (Takahashi and Yamanaka, 2006; Yu et al., 2007) are only a few of a great number of step stones that completely changed our understanding of stem cells and brought stem cell based therapy from the field of science fiction into reality (Hernández et al., 2011; Brunt et al., 2012; Bernstein, 2012; Kanno, 2013; Lin et al., 2013; Okano et al., 2013). All the recent achievements in stem cell biology also made clear the technical limitations that scientists are facing every day in their quest to unriddle the secrets of stem cells, limitations that become harder and harder to overcome. One of the key technical problems in the field today (and we discussed it in detail in the Introduction part of this thesis) is a shortage of models that would allow to study pluripotent stem cell *in vivo*, within its niche and its physiological context. It is hard to overestimate the importance of traditional model systems, such as mice, fishes or flies, for the stem cell research, yet the need for new complementary models becomes obvious (Sánchez Alvarado, 2004; Gurley et al., 2008) .

Understanding of this problem made scientists search for the new model organisms that could be used to expand our knowledge of stem cell biology. In the recent years a handful of such models are given momentum as new promising tools in the field. Sea anemone *Nematostella vectensis* (Darling et al., 2005; Watanabe et al., 2009) and a freshwater cnidarian *Hydra* (Bosch et al., 2010; Galliot, 2012), various ascidian (Laird et al., 2005; Tiozzo et al., 2008) and flatworm species (Aboobaker, 2011; Gentile et al., 2011; Rink, 2013) are the great illustrations of this newly emerged campaign for establishing new or expanding the capacities of already known model organisms for *in vivo* stem cell research.

Our group works on developing the marine free-living flatworm *Macrostomum lignano* (Ladurner et al., 2005) as a new suitable model for stem cell and regeneration studies. *M. lignano* possess a population of adult pluripotent stem cells called neoblasts; it is easy to culture and work with in the laboratory conditions, and the impressive toolbox to work with this worm was developed in the last 10 years (Nimeth et al., 2004; Ladurner et al., 2005; Pfister et al., 2007; Pfister et al., 2008; De Mulder et al., 2009; De Mulder et al., 2010; Demircan and Berezikov, 2013). All these features, however, are shared between *Macrostomum* and well-known planarian flatworm models (Gentile et al., 2011). What really does make our model special is its amenability to genetic approaches and availability of transgenics, recently developed by our group (Demircan et al., in preparation). Another advantage of *M. lignano* is its basal position on the evolutionary tree of the flatworms and, therefore, of all bilaterians that gives us a unique opportunity to look into the evolutionary aspects of stem cell regulation (Ehlers, 1985; Riutort et al., 2012).

The main aim of this thesis was to explore genomic resources of *M. lignano* and to demonstrate the potential of this organism as a model system for stem cell research and the fields beyond. The results of this work are briefly discussed below.

1. ADVANTAGE OF BEING A BASAL FLATWORM

M. lignano represents a basal member of free-living worms within the Rhabditophora, the largest taxon in the phylum Platyhelminthes, and thus can be considered one of the most basal bilaterian species. This position makes it extremely interesting to study *Macrostomum* from the evolutionary point of view.

In Chapter 2 we presented initial results for *de novo* sequencing of the genomes of *M. lignano* and two other species - very close-related *Macrostomum hystrix* and the acoel *Isodiametra pulchra*. Sequenced and annotated genome is a necessary tool for any further studies in the model organism, and the assembly we are using at the moment is already good enough for the discovery of most of the genes expressed in *M. lignano*. At the same time the genome data from *Macrostomum* gives us a unique opportunity to look into evolution of pluripotency. Comparison of gene set of *M. lignano* with recent data from *N. vectensis* (Putnam et al., 2007) would provide a great insight into the genome composition of common eumetazoan ancestor and the conservation of the mechanisms involved in stem cell regulation. Additionally, incorporation of *M. hystrix* data into the same analysis allows estimation of the importance of genus-specific genes, only conserved among close-related species, in the same processes.

Position of Acoels within the evolutionary tree is currently under debate. From one point of view they are considered to be a group of very basal bilaterians, more primitive than flatworms (Egger et al., 2009). Alternatively, Acoels could be placed much closer to chordates (Philippe et al., 2011). In both cases comparative analysis of the genomes of *M. lignano* and *I. pulchra* is very interesting, for it would rather help to reconstruct the genome of the common bilaterian ancestor and identify the set of genes potentially responsible for the emergence of bilaterally symmetrical body plan, or to track the evolution of pluripotency from basal flatworms to much more evolutionary derived species.

The *M. lignano* genome project is still ongoing and is not ready yet for a thorough evolutionary analysis mentioned above, which is also beyond the scope of this Thesis. At the same time, the preliminary results from the *M. lignano* genome initiative were already extremely useful for a number of other studies in this model organism and in development of transgenesis methods in particular. While the repetitive nature of *M. lignano* genome, combined with potential partial genome duplication, complicate genome assembly substantially, it is envisioned that recently acquired additional data from different Next Generation sequencing platforms would improve the currently used genome assembly greatly in the near future.

In Chapter 6 we discussed miRNAs expressed in *M. lignano* and present a set of miRNAs potentially enriched in neoblasts that might play an important role in stem cell functioning. This is the first study describing miRNAs in *M. lignano*, and it also provides the first data on differential expression of certain miRNA species in irradiated and non-irradiated *Macrostomum* worms and presents the reliable method to identify expression patterns of miRNAs in this model organism. What might be even more interesting, in this chapter we describe highly unusual conformation of pre-miRNA hairpins in *M. lignano*, which, given the evolutionary position of the species, could provide a great insight into the emergence and evolution of miRNAs in animals – the process that still remains largely unclear. We are currently working on sequencing of miRNAs of acoel *Isodiametra pulchra*, and this new data would potentially add greatly to our findings.

Thus, genome and miRNA sequencing data acquired from *M. lignano* not only provide the necessary tool for further research in this model organism, but also raise important evolutionary questions and set a direction for interesting studies.

2. HUNT FOR NEOBLAST-SPECIFIC GENES

In Chapter 3 we presented first transcriptome assembly for *M. lignano*. Just as in cases of genome and miRNA profiling, simply having this data is extremely important for any future studies in this organism, for it is virtually impossible in modern biological science to perform any genetic or molecular research in the model without having full information about the set of genes expressed in it. Importantly, this assembly provides a reference for sets of RNA-seq data (see also part 3 of this General Discussion) that can be acquired from animals (or *M. lignano* cell populations) at different conditions. The assembly we made appears redundant, but it is also complete, covering the absolute majority of the genes expressed in *M. lignano*.

In the same chapter we used 3 independent RNA-seq based approaches to predict set of genes with potentially enriched expression in neoblasts of *M. lignano*. RNA-seq reads were mapped to the original transcriptome assembly, proving its reliability and usefulness for this approach. Two of the three approaches were based on the comparison of expression data from irradiated and control (mock-irradiated worms). Elimination of dividing cells of flatworms by irradiation has long been used as experimental approach to find neoblast-specific transcripts (Salvetti et al., 2000; Hayashi et al., 2006; Eisenhoffer et al., 2008; Solana et al., 2012). In the first approach we simply compared expression data from irradiated and intact worms and looked at the genes that would be most significantly down-regulated following irradiation. Another approach was based on expression dynamics of the genes following irradiation, as we assumed that neoblast specific genes should demonstrate consistent decline in expression after the exposure to the irradiation dose. In the third approach we compared RNA-seq data from adult and juvenile (lacking gonads) worms in order to identify genes enriched specifically in gonads and exclude them from our screen.

Based on these approaches we generated three different candidate lists, and the one based on consistent decline of expression after irradiation proved to be the most reliable to predict neoblast-specific genes. At the same time, expression of the majority of the candidates (analyzed by *in situ* hybridization) was restricted to proliferative compartments of the worm body. This proves generally that quantitative RNA-seq could be a useful tool for gene expression pattern and function prediction. And combining different sets of RNA-seq data would help to understand functions and importance of given gene in more detail and define better criteria to look for genes specific for certain cell population. Currently our group is working on expression profiling of sorted neoblasts and generating RNA-seq data from regenerating animals and, more specifically, isolated anterior and posterior blastemas is on the list of our future plans. Combination of all these expression data sets would provide a useful tool for studying regeneration, development and stem cell functioning in general.

One of the important advantages of *M. lignano* as a model for such screens is the fact that two populations of virtually pluripotent cells, namely neoblasts and germ line cells are present in its body side-by-side. There is a large overlap between the sets of genes expressed

in these two populations (Shibata et al., 2010; Wang et al., 2010), and our results confirmed it once more, yet the actual relations (evolutionary and functional) between neoblasts and germ cells remain largely unknown. *M. lignano* provides a powerful model system, where two lineages of pluripotent stem cells, their embryonic origins and potential differences in mechanisms that establish and maintain pluripotency in both populations could be studied at the same time.

Also in Chapter 3 we addressed functions of a number of candidate neoblast-specific genes by RNA interference and described 9 stem cell associated knockdown phenotypes. We described *DDX39* as a novel highly conserved stem cell specific gene, playing a crucial role in cell proliferation, not just in *M. lignano*, but also in human cell lines.

3. SAME GENE, DIFFERENT POINTS OF VIEW

Another study, based on RNA-seq approach is presented in Chapter 5. We described the effect of extremely low salinity conditions on the reproductive system of *M. lignano*, and addressed the molecular mechanism underlying this effect by comparing expression data from the worms exposed to low salinity and control animals over a time course. We identified groups of genes potentially involved in different physiological processes such as stress-response, starvation-related metabolism changes, salinity tolerance and, most importantly, rapid gonad degradation caused by low salinity conditions.

During the studies described in Chapters 3 and 5 we collected a unique set of transcriptome-wide expression data. Our RNA-seq data give us an opportunity to look at the expression dynamics of every gene in *M. lignano* at different developmental stages and under variety of conditions. Other sets of data we generated that were not presented in this thesis include RNA-seq of the worms at 4 different stages of embryonic development. We also briefly mentioned above the plans to obtain expression data from sorted neoblasts, regenerating animals and isolated blastemas. Continuation of the project presented in Chapter 5 would include sequencing the transcriptome of worms exposed to high salinity. There are also RNA-seq data sets generated by other groups working on *M. lignano*, including the expression profiling of different fragments of the worm body along the anterior-posterior axis (Arbore et al., in preparation). Finally, advances in generating transgenic lines push us towards the idea of expression profiling of worms carrying a knockout or over-expressing a certain gene in order to address its importance and function in a great detail.

With the costs of transcriptome-wide RNA-seq experiments going down every year we have a unique opportunity to collect extensive information on the expression of every gene in *M. lignano*. Our group is currently working on establishing an online open source database (http://www.macgenome.org), where we plan to gather all possible data about *M. lignano* transcripts, including expression levels at different developmental stages, under different conditions and within different isolated cell populations, expression patterns and RNAi phenotypes. Having such a tool available would greatly improve our knowledge of any given gene function in the variety of biological processes.

4. GOING APPLIED: DRUG AND TOXIN SCREENS

Studies of bioelectric signals and their effects on stem cell functioning recently emerged as a completely new field of stem cell biology. Ion flows and membrane voltage were shown to control cell proliferation (Blackiston et al., 2009) and differentiation (Sundelacruz et al., 2008) or, on the level of the organism, morphogenesis during development, regeneration and cancer progression (Adams, 2008; Levin, 2009; Levin, 2012). In Chapter 4 we discussed the potential of *M. lignano* as a model for studies of bioelectric signaling. We showed that membrane voltage pattern of marine *M. lignano* on the scale of the whole body is similar to the one observed in freshwater *Dugesia japonica* (Beane et al., 2011). Re-establishing of this pattern in regenerating worms (specifically, in the posterior fragments), however, happens very quickly in *Dugesia* but never occurs in *Macrostomum*, which could explain the lack of regeneration in anterior-facing wounds of *M. lignano*. We showed the results of pilot experiments, demonstrating that changing membrane voltage in regenerating worms (by applying water-soluble drugs to the media) can greatly change the regenerating potential of blastemas.

Main conclusion of this study is that *M. lignano* can be successfully used as a model for bioelectric signaling research. It is easy to apply drugs that affect membrane voltage and dyes that help to visualize membrane potentials and ion concentrations to the worms, and, importantly, to screen the worms for stem cell associated phenotypes. Membrane voltage gradients in the body are established and maintained by the orchestrated functioning of many ion channels and transporters, and in our screen for genes, expression of which is enriched in neoblasts (Chapter 3), we find a number of ion channel coding genes. Further studies of these candidates could provide an insight into the link between membrane potentials and stem cell functioning.

Our work also demonstrated the advantage of *M. lignano* as a model reporter organism for virtually any drug or toxin screens. We already mentioned that it is easy to apply any agents to this water-living animal. Transparency and clear morphology help to read and interpret physiological effects, while small size, short generation time and efficient reproduction make logistics of large-scale experiments perfectly feasible. Importantly, availability of transgenics offers an opportunity to create unique reporter lines that would better suit certain experiment conditions. Developing of automated screening techniques would add greatly to the advantages of *M. lignano* as a model for large-scale research. In conclusion, we believe that *M. lignano* has a bright future as a model not only for bioelectrical and ion channel studies, but also for any drug and toxin screens, especially for those concentrated on stem cell related readout.

CONCLUDING REMARKS

4 years ago, when our group just started working with *M. lignano*, there were only two PhD students in the laboratory involved in research on this model organism. First worm cultures in our laboratory originated from the groups of Peter Ladurner and Lukas Schärer who were working with *M. lignano* since late 1990-s and early 2000-s respectively. Together with the animals we took from them a developed toolkit that already included ISH, RNAi and BrdU-labeling techniques, irradiation method, set of tissue-specific antibodies and, importantly, robust culturing protocols (Nimeth et al., 2004; Ladurner et al., 2005; Pfister et al., 2007;

Pfister et al., 2008; De Mulder et al., 2009; De Mulder et al., 2010). The aim of this thesis was to further develop *M. lignano* as a reliable model system for stem cell research, what in part meant expanding this toolkit. And we added knowledge of the genomic resources of *M. lignano* to it. By that we mean, first of all, *de novo* sequencing, assembling and annotation of genome and transcriptome and miRNAs profiling, but also a massive data on expression of all the genes in *M. lignano* at different developmental stages and under different conditions. Today most of this data are publicly available, and with the introduction (in a near future) of the new online database on *M. lignano* gene expression everybody interested in this model would have an opportunity to benefit from the results of our work. Another methodological outcome of this thesis is the introduction of reliable miRNA ISH method, suitable for *M. lignano*.

In a course of this thesis we did not only improve the toolkit, but also tested the potential of *M. lignano,* as a model for different kinds of research and studies. Our focus was always on the stem cell related topics, but it is not the only field where *M. lignano* could appear useful. First, it is an extremely interesting organism to look at from the evolutionary point of view, and both its genome and its presumably ancient miRNA system could add a lot to our understanding of evolution of many physiological processes. Second, we showed that *M. lignano* can serve as a useful model for various chemical compound (toxin or drug) screens. Studying of bioelectricity and ion channels, although connected to stem cell research, presents a third field where many advantages of this worm model can be exploited. It is worth mentioning that today *M. lignano* is already successfully used as a model for the whole array of studies, including sex allocation (Schärer et al., 2005; Janicke and Schärer, 2010; Janicke et al., 2013) and ageing (Mouton et al., 2009).

Finally, our studies also brought us one step forward towards understanding the functioning of stem cells in *M. lignano*. We identified a set of genes involved in stem cell regulation during homeostasis and regeneration, and even larger group of transcripts, expression of which is enriched in neoblasts. We also made a candidate list of miRNAs potentially involved in stem cell regulation and addressed the role of ion channels and bioelectric signaling in regeneration. Finally, we initiated the studies of the mechanism involved in rapid gonad degradation following exposure to the low salinity and presented the first results of this project. Some of the projects we were working on in parallel to the ones described in this thesis (for example, studies of the canonical Wnt signaling in *M. lignano*) were left out of this book, but their results would certainly complement the ones described above and add to our general understanding of stem cells in *M. lignano*.

So these are three parts of the results of this thesis as we see them: developing new tools, testing new fields and bringing up first solid results. All three serve the purpose to develop *M. lignano* as a reliable model organism for stem cell research, but even more importantly, they set directions for numerous future studies – directions that we hope the others would follow, as we would ourselves. Since the work described in this thesis was initiated, the number of people in our laboratory working with *M. lignano* increased greatly, as did the number of groups in the world using this model in their studies. We hope that the work presented in this book played its role in attracting a lot of interest to this fairly new model organism. And we also hope that the growth of the *Macrostomum* community will continue for it would take the efforts of many people to fully explore all the benefits this model system provides.

REFERENCES

Aboobaker, A.A. (2011). Planarian stem cells: a simple paradigm for regeneration. Trends Cell Biol *21*, 304-311.

Adams, D.S. (2008). A new tool for tissue engineers: ions as regulators of morphogenesis during development and regeneration. Tissue Eng Part A *14*, 1461-68.

Beane, W.S., Morokuma, J., Adams, D.S., and Levin, M. (2011). A chemical genetics approach reveals H,K-ATPase-mediated membrane voltage is required for planarian head regeneration. Chem Biol *18*, 77-89.

Bernstein, H.S. (2012). Cardiac repair and restoration using human embryonic stem cells. Regen Med *7*, 697-712.

Blackiston, D.J., McLaughlin, K.A., and Levin, M. (2009). Bioelectric controls of cell proliferation: ion channels, membrane voltage and the cell cycle. Cell Cycle *8*, 3519-528.

Bosch, T.C., Anton-Erxleben, F., Hemmrich, G., and Khalturin, K. (2010). The Hydra polyp: nothing but an active stem cell community. Dev Growth Differ *52*, 15-25.

Brockes, J.P., and Kumar, A. (2008). Comparative aspects of animal regeneration. Annu Rev Cell Dev Biol *24*, 525-549.

Brunt, K.R., Weisel, R.D., and Li, R.K. (2012). Stem cells and regenerative medicine - future perspectives. Can J Physiol Pharmacol *90*, 327-335.

Darling, J.A., Reitzel, A.R., Burton, P.M., Mazza, M.E., Ryan, J.F., Sullivan, J.C., and Finnerty, J.R. (2005). Rising starlet: the starlet sea anemone, Nematostella vectensis. Bioessays *27*, 211-221.

Demircan, T., and Berezikov, E. (2013). The Hippo pathway regulates stem cells during homeostasis and regeneration of the flatworm Macrostomum lignano. Stem Cells Dev *22*, 2174-185.

Egger, B., Steinke, D., Tarui, H., De Mulder, K., Arendt, D., Borgonie, G., Funayama, N., Gschwentner, R., Hartenstein, V., et al. (2009). To be or not to be a flatworm: the acoel controversy. PLoS One *4*, e5502.

Ehlers, U. (1985). phylogenetische System der Plathelminthes (G. Fischer).

Eisenhoffer, G.T., Kang, H., and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian Schmidtea mediterranea. Cell Stem Cell *3*, 327-339.

Galliot, B. (2012). Hydra, a fruitful model system for 270 years. Int J Dev Biol *56*, 411-423.

Gentile, L., Cebrià, F., and Bartscherer, K. (2011). The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration. Dis Model Mech *4*, 12-19.

Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008). Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science *319*, 323-27.

Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Dev Growth Differ *48*, 371-380.

Hernández, J., Torres-Espín, A., and Navarro, X. (2011). Adult stem cell transplants for spinal cord injury repair: current state in preclinical research. Curr Stem Cell Res Ther *6*, 273-287.

Janicke, T., and Schärer, L. (2010). Sperm competition affects sex allocation but not sperm morphology in a flatworm. Behavioral Ecology and Sociobiology *64*, 1367-375.

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D.B., and Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution *67*, 3233-242.

Kanno, H. (2013). Regenerative therapy for neuronal diseases with transplantation of somatic stem cells. World J Stem Cells *5*, 163-171.

Kiger, A.A., White-Cooper, H., and Fuller, M.T. (2000). Somatic support cells restrict germline stem cell self-renewal and promote differentiation. Nature *407*, 750-54.

King, R.S., and Newmark, P.A. (2012). The cell biology of regeneration. J Cell Biol *196*, 553-562.

Ladurner, P., Pfister, D., Seifarth, C., Schärer, L., Mahlknecht, M., Salvenmoser, W., Gerth, R., Marx, F., and Rieger, R. (2005). Production and characterisation of cell- and tissue-specific monoclonal antibodies for the flatworm Macrostomum sp. Histochem Cell Biol *123*, 89-104.

Ladurner, P., Schärer, L., Salvenmoser, W., and Rieger, R.M. (2005). A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: Macrostomum lignano, n. sp.(Rhabditophora, Macrostomorpha). Journal of Zoological Systematics and Evolutionary Research *43*, 114-126.

Laird, D.J., De Tomaso, A.W., and Weissman, I.L. (2005). Stem cells are units of natural selection in a colonial ascidian. Cell *123*, 1351-360.

Levin, M. (2009). Bioelectric mechanisms in regeneration: Unique aspects and future perspectives. Semin Cell Dev Biol *20*, 543-556.

Levin, M. (2012). Morphogenetic fields in embryogenesis, regeneration, and cancer: non-local control of complex patterning. Biosystems *109*, 243-261.

Lin, H.T., Otsu, M., and Nakauchi, H. (2013). Stem cell therapy: an exercise in patience and prudence. Philos Trans R Soc Lond B Biol Sci *368*, 20110334.

Mouton, S., Willems, M., Braeckman, B.P., Egger, B., Ladurner, P., Schärer, L., and Borgonie, G. (2009). The free-living flatworm Macrostomum lignano: a new model organism for ageing research. Exp Gerontol *44*, 243-49.

De Mulder, K., Kuales, G., Pfister, D., Egger, B., Seppi, T., Eichberger, P., Borgonie, G., and Ladurner, P. (2010). Potential of Macrostomum lignano to recover from gamma-ray irradiation. Cell Tissue Res *339*, 527-542.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., et al. (2009). Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. Dev Biol *334*, 198-212.

Nimeth, K.T., Mahlknecht, M., Mezzanato, A., Peter, R., Rieger, R., and Ladurner, P. (2004). Stem cell dynamics during growth, feeding, and starvation in the basal flatworm Macrostomum sp. (Platyhelminthes). Dev Dyn *230*, 91-99.

Okano, H., Nakamura, M., Yoshida, K., Okada, Y., Tsuji, O., Nori, S., Ikeda, E., Yamanaka, S., and Miura, K. (2013). Steps toward safe cell therapy using induced pluripotent stem cells. Circ Res *112*, 523-533.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in Macrostomum lignano. Dev Biol *319*, 146-159.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. (2007). The exceptional stem cell system of Macrostomum lignano: screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. Front Zool *4*, 9.

Philippe, H., Brinkmann, H., Copley, R.R., Moroz, L.L., Nakano, H., Poustka, A.J., Wallberg, A., Peterson, K.J., and Telford, M.J. (2011). Acoelomorph flatworms are deuterostomes related to Xenoturbella. Nature *470*, 255-58.

Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., et al. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science *317*, 86-94.

Rink, J.C. (2013). Stem cell systems and regeneration in planaria. Dev Genes Evol *223*, 67-84.

Riutort, M., Álvarez-Presas, M., Lázaro, E., Solà, E., and Paps, J. (2012). Evolutionary history of the Tricladida and the Platyhelminthes: an up-to-date phylogenetic and systematic account. Int. J. Dev. Biol *56*, 5-17.

Salvetti, A., Rossi, L., Deri, P., and Batistoni, R. (2000). An MCM2-related gene is expressed in proliferating cells of intact and regenerating planarians. Dev Dyn *218*, 603-614.

Sánchez Alvarado, A. (2004). Regeneration and the need for simpler model organisms. Philos Trans R Soc Lond B Biol Sci *359*, 759-763.

Schärer, L., Sandner, P., and Michiels, N.K. (2005). Trade-off between male and female allocation in the simultaneously hermaphroditic flatworm Macrostomum sp. J Evol Biol *18*, 396-404.

Shibata, N., Rouhana, L., and Agata, K. (2010). Cellular and molecular dissection of pluripotent adult somatic stem cells in planarians. Dev Growth Differ *52*, 27-41.

Solana, J., Kao, D., Mihaylova, Y., Jaber-Hijazi, F., Malla, S., Wilson, R., and Aboobaker, A. (2012). Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNA-seq, RNA interference and irradiation approach. Genome Biol *13*, R19.

Sundelacruz, S., Levin, M., and Kaplan, D.L. (2008). Membrane potential controls adipogenic and osteogenic differentiation of mesenchymal stem cells. PLoS One *3*, e3737.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tiozzo, S., Brown, F.D., De Tomaso, A.W. (2008) Regeneration and stem cells in ascidians. In Stem Cells, 95-112. Springer.

Wang, Y., Stary, J.M., Wilhelm, J.E., and Newmark, P.A. (2010). A functional genomic screen in planarians identifies novel regulators of germ cell development. Genes Dev *24*, 2081-092.

Watanabe, H., Hoang, V.T., Mättner, R., and Holstein, T.W. (2009). Immortality and the base of multicellular life: Lessons from cnidarian stem cells. Semin Cell Dev Biol *20*, 1114-125.

Xie, T., and Spradling, A.C. (2000). A niche maintaining germ line stem cells in the Drosophila ovary. Science *290*, 328-330.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-920.

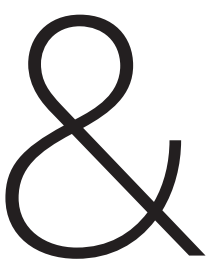
APPENDIX

NEDERLANDSE SAMENVATTING

ACKNOWLEDGEMENTS

CURRICULUM VITAE

PUBLICATIONS



NEDERLANDSE SAMENVATTING

De afgelopen twee decennia worden gekenmerkt door ontdekkingen en doorbraken in verschillende disciplines in de biologie. Daarbij is de stamcelbiologie een voorbeeld van een snel ontwikkelend veld. Ontdekking van stamcelniches en het herprogrammeren van somatische cellen naar een pluripotente toestand zijn slechts enkele voorbeelden van ontdekkingen die onze kijk op stamcellen hebben veranderd. Door alle recente doorbraken in de stamcel biologie zijn echter ook de technische beperkingen waar wetenschappers dagelijks tegenaan lopen duidelijk geworden, beperkingen die steeds duidelijker aan het licht komen. Het gebrek aan modelsystemen waarin pluripotente stamcellen *in vivo* in hun eigen micromillieu kunnen worden bestudeerd vormt een van de belangrijkste technische tekortkomingen. Het is niet moeilijk om je voor te stellen hoe belangrijk traditionele modelsystemen als muizen, vissen en vliegen zijn. Echter, de noodzaak voor nieuwe aanvullende modelsystemen wordt steeds duidelijker. In **Hoofdstuk 1** van dit proefschrift hebben we de huidige ontwikkelingen in de moderne stamcelbiologie bediscussieerd. We hebben ons toegelegd op de modelsystemen die op dit moment worden gebruikt in het veld en hun bijbehorende stamcelsystemen. We hebben ook de veelbelovende stamcelmodellen in de platworm besproken en hebben de *Macrostomum lignano* als nieuw stamcelmodelsysteem geïntroduceerd. De laatste zal een aanvulling worden op bestaande planaria modellen.

De platworm *M. lignano* is recentelijk als modelorganismse naar voren gekomen. Het bezit een populatie van adulte somatische stamcellen, neoblasten genaamd, die het tot een aantrekkelijk modelsysteem maakt voor stamcelstudies. Het ontwikkelen van genomische hulpbronnen en, meer algemeen, het tot een betrouwbaar model ontwikkelen van *M. lignano* zijn de belangrijkste doelen in dit proefschrift. We benaderen deze doelen door de instrumenten binnen het model uit te breiden, de bruikbaarheid van het model te toetsen in verschillende wetenschappelijke contexten en het functioneren van het stamcelsysteem te bestuderen in het organisme.

Aan de bestaande instrumenten die we voor *M. lignano* ter beschikking hebben, hebben we de kennis van zijn genomische hulpbronnen toegevoegd. Daarmee bedoelen we *de novo* sequencing, samenvoeging en annotatie van het genoom en transcriptoom (beschreven in respectievelijk **Hoofdstuk 2** en **3**) en miRNA profiling (**Hoofdstuk 6**), maar ook het genereren van kennis over gen expressie in *M. lignano* tijdens verschillende ontwikkelingsstadia en condities. Het meerendeel van deze data, beschreven in **Hoofdstuk 3** en **5** is al openlijk toegankelijk en in de nabije toekomst introduceren we een nieuwe database met daarin *M. lignano* genexpressie data waardoor iedereen van onze bevindingen gebruik kan maken. Aan de lijst van methoden om *M. lignano* te bestuderen hebben we ook een betrouwbare mRNA *in situ* hybridizatie methode toegevoegd (beschreven in **Hoofdstuk 6**).

We hebben ook de mogelijkhied om *M. lignano* als model te gebruiken voor verschillende studies buiten het stamcelveld getoetst. Ten eerste is het vanuit een evolutionair oogpunt een interessant organisme, en zowel zijn genoom (**Hoofdstuk 2**) als het vermoedelijk zeer oude miRNA systeem (besproken in **Hoofdstuk 6**) zouden veel toe kunnen voegen aan onze kennis over de evolutie van een verscheidenheid aan fysiologische processen. Ten tweede: *M. lignano*

kan dienen als een model voor screens naar chemische agentia. Het bestuderen van bioelectrische processen vormt een derde mogelijkheid waarvoor dit wormmodel gebruikt kan worden. De voordelen van *M. lignano* als model voor de laatste twee onderzoeksvelden worden bediscussieerd in **Hoofdstuk 4**.

We hebben ook een stap voorwaards gemaakt in het begrip van het functioneren van stamcellen in *M. lignano*. We hebben een set genen geïdentificeerd die betrokken zijn bij de regulatie van stamcellen tijdens homeostase en regeneratie, en een nog grotere groep transcripten waarvan de expressie in neoblasten is toegenomen (zie **Hoofdstuk 3**). In **Hoofdstuk 6** hebben we een kandidatenlijst gemaakt van miRNAs die mogelijk zijn betrokken bij de regulatie van stamcellen en hebben we een rol toegeschreven aan bioelectrische signalering in regeneratie in **Hoofdstuk 4**. Tot slot zijn we van start gegaan met de studies naar de mechanismen betrokken bij de snelle degradatie van de gonaden als gevolg van blootstelling van een lage zoutspiegel. De eerste resultaten van dit werk worden in **Hoodstuk 5** gepresenteerd.

In **Hoofdstuk 7** worden de belangrijkste resultaten en conclusies van de studies in dit proefschrift samengevat. In dit hoofdstuk bediscussieren we niet enkel hetgeen wat bereikt is in de afgelopen vijf jaar, maar blikken we ook vooruit op mogelijke toekomstige studies.

ACKNOWLEDGEMENTS

This thesis comes to an end, and so does one of the most interesting periods of my life so far. There are many people without whom the completion of this work would have never been possible, and even more (way more!) without whom it could have been much less fun. Simply putting here the names of all the people I want to thank (let alone saying a couple of personal words to each of them) would make this book twice the size. Yet I will take the risk and mention at least some of them. I will try to make it short, and of course there will be some great people who would not find themselves in the section below. And I apologize to them in advance.

First of all, I want to express my gratitude to my boss. **Zhenya**, thank you for taking me to your team. 5,5 years ago I was looking for the PhD position in the Netherlands, and I also wanted to work on regeneration, preferably with an interesting model organism. It did not take long for me to give up to find such a group and start looking into different topics. I remember, how I felt, when you mentioned during the interview, that you plan to start working with a new flatworm model organism that can regenerate really well. I think it was fate, and we both simply accepted it. I learned a lot from you, and working under your guidance was never boring. And thank you for letting me stay in the lab for another year - that meant a lot to me. It is really great to see the lab grow and do well at the new place. I wish you all the best in your scientific career, and let the science never become boring to you!

I also wanted to thank all the **members of my reading committee** for their patience and understanding. And for the final approval of this thesis for the defense, despite all the well-deserved criticism. And special thanks to **Rik** personally - for being a chair of the reading committee and for all the time we shared and Hubrecht and all the things I have learned from you. I still remember - it was one of my first weeks in Hubrecht, and you showed me how to seal the chamber for agarose gel with the tape - since then my gels never leaked!

My dear labmates! It was an honor, a pleasure and a great fun to work together with you. I learned something from each of you! **Meltem**, you were the very first person working in Zhenya's lab, and I was the second. Thank you for introducing to me to Hubrecht and the routine in the laboratory. **Turan**, it was great to share these 4,5 years in the lab with person so dedicated to science. We didn't have as many discussions as we probably wanted, but all the talks we had were very fruitful and interesting, and we always had support and understanding of each other, and that was very important for me. **Katrien,** you changed the way I looked at many things - not only in science. I would remember the talks we had in the lab (sometimes well after midnight), your trust in me and my work, and your very positive look at life and everything around you. Thank you! Dear students - **Lin**, **Imre**, **Bowo**, **Kay**! You were all very different, but I liked and respected each of you! I was happy that Lin and Imre stayed in the lab for longer, as technicians - you were a great help and support to the work described in this thesis and beyond the scope of it! You all chose very different paths - family, travelling, science and business, and I wish you all the best and a great success in whatever you are doing now. **Stijn**, you taught me how important it is in science to pay attention to the details. Thank you for sharing your experience with me and for all the comments you gave me while I was writing some parts of this thesis. I was stubborn enough not to follow some of them, but it does not mean that I did not respect and value your input. **Magda**, you are probably the best person in

our lab to put science and fun together, and I always liked this attitude very much. Besides, you are a good scientist, and I apologize if I ever made you feel otherwise! Finally, I also want to think all the "new" Berezikov group members, who never worked in Hubrecht, but joined the lab in Groningen. **Lisa**, **Jakub**, **Philipp**, **Katya**, **Evgenia**, **Daniel**, **Frank**, **Margriet**, **Nikos** - it is great to join you guys, and I am happy that we have a bit of time ahead to share and get to know each other better!

There were two people who introduced me to the beautiful model I fell in love with and spent most of my PhD exploring, two "parents" of Macrostomum community. **Peter** and **Lukas**, thank you for bringing this funny creature into my life, and for all the knowledge you were always ready to share. Lukas and **Dita**, I owe a special thanks to you for your hospitality. Hosting Macrostomum meetings, you made me feel at home in Basel!

Among all the groups in Hubrecht Institute there is one, that have always been special - in fact very special for me personally. Cuppen pack, I never had a chance to express it properly, but I owe you a lot. **Pim**, **Henk**, **Bas** (not really from the Cuppen lab, but always somewhere around), **Ewart**, **Mark**, **Joram**, **Frans Paul** and a bit later - **Nico** - you were the first people to introduce me to the social life not only in Hubrecht, but in the whole Netherlands. I would never forget that. You made me feel comfortable in this country and in the Institute. Often being the only non-Dutch person in the Rex, I never felt an alien, not a little. I still remember my first birthday in the Netherlands - spent with you; the first bowling experience in the corridors of Hubrecht; first Tequila Reverse (well, I better forget this one)... People in Russia sometimes say that it must be boring - living and working in the Netherlands. Well, they never met you. And, of course, **Edwin**! Thank you for being my promoter and for setting a great example of a group leader.

In 2009 I had a great luck of joining the Hubrecht PV. This responsibility I shared with the best women around. Girls, you already know how I feel about you, and I can never thank you enough for all the time we spent together. **Flore**, **Maartje**, **Paula**, **Petra**, **Tamara** - rarely a man has an opportunity to be surrounded by such different yet equally beautiful, smart and kind girls, and I was the lucky bastard! Thank you for all the time we spent together and I am sure we will keep in touch!

Leon and **Teije**, you agreed to be my paranymphs at the defense, and I consider it an honor. And not only because you are both taller than me, and look great behind my back. Teije, you are probably the most unusual, the most "different" Dutch person I ever met. Yes, you still love cycling and beer, but that would probably cover the typically Dutch features of your character. It was a great experience to have discussions with you regardless of the topic, and your opinion was always very interesting for me. Leon, you introduced me to the world of Dutch football turning me (and my wife) into FC Utrecht fans. You also organized the wonderful trip to Mallorca that was a lot of fun. But what is more important you showed me how it is to have a Dutch friend - thank you for that!

This brings me to you, **Sjoerd**. I am lucky and honored to call you my friend, and, honestly, there is not much else to say, for the words, especially written would never express enough. I hope to visit you and Tamara soon in Switzerland, and my home (would it be in the Netherlands, Russia or elsewhere) is always opened for you - but I guess you know it already. And big thanks

to all the fellow de Laat's people I got to know - **Yuva**, **Yun**, **Patrick** and the others. We didn't spend much time together, especially lately, but whenever we did, it was always a great fun!

A couple of words to Korswagen guys and van Rheenen girls! There are rooms in the Institute that you only enter for a couple of seconds to quickly ask somebody a question or to get a person you want to speak with outside. **Anoek**, **Evelyne**, **Saskia**, **Reinoud**, **Remco**, **Marco** - whenever I would come to one of your offices I always felt welcome and wanted to stay and talk. I am sorry for all the moments when it actually disturbed you.

People I shared the lab space with - **Ruben**, **Jos**, and later - **Charles** and **Maartje**. Thank you for turning work into a great fun! Dear office-mates, **Magdalena**, **Martin**, and **Robin people** - I want to thank you as well!

All the people from the **third floor** I did not mention before - it was a pleasure to have you around! Thank you very much!

People who are usually around on Friday afternoon (strangely enough the same people could often be found in the lab on weekend) - I really enjoyed all the moments we shared. **Alex**, **Manda**, **PJ**, **Mauro**, **Eirinn** - I would not know what to do on Friday afternoons if it was not for you. **Maaike**, having a wonderful girl living on top of the great pub is probably every man's dream! I hope to see you around Ledig Erf sometimes! **Emily** and **Paul** - I loved our Monday afternoon beers!

I also want to thank all my **sailing friends** for the great effort they made to ruin my PhD and prevent this thesis defense from happening! You really did well and you were wise and kind enough not to do it too well.

The last people I want to mention here, though they definitely should be the first in order of importance, are my family. I want to thank **my parents** for having faith in me and showing support at the hard times. I hope I lived up to your hopes, and I know that I would never make it this far without your help and your love. And **my wife** - the one and only reason I do more or less everything I do. All these years you believed in me more than I ever believed in myself - thank you!

CURRICULUM VITAE

Dan Simanov was born on the 11th of May 1986 in Moscow, Russia. In 2003 he finished secondary education *cum laude* at the Moscow Gymnasium 1543 and entered the Biological Department of Lomonosov Moscow State University. During the studies he performed a long internship (2 years) at the Moscow Institute of Carcinogenesis, N.N. Blokhin Cancer Research Center, where he studied the role of hepatocyte nuclear factor 4 alpha (HNF4α) in hepatocellular carcinoma progression under supervision of Natalia Lazarevich. In 2006 and 2007 he also had two internships (3,5 months each) in Max Planck Institute of Molecular Cell Biology and Genetics in Dresden, working in the group of Elly Tanaka. There Dan's focus was on introducing sequences for site-specific recombination into the genome of axolotl transgenic lines and establishing the protocol for confocal microscopy-based analysis of the neural crest regulation in axolotls. He completed his Mater studies in 2008 and obtained the position in the Berezikov Lab at the Hubrecht Institute in Utrecht. Part of the work performed in this group is presented in this book.

PUBLICATIONS

Simanov, D., Mellaart-Straver, I., Sormacheva, I., and Berezikov, E. (2012). The flatworm *Macrostomum lignano* is a powerful model organism for ion channel and stem cell research. Stem Cells Int *2012*, 167265.

Simanov, D., Wasik K.A., Canela A., Mouton S., van Zon, P., de Mulder, K., de Bruijn, E., Cuppen, E., Aboobaker, A., Qi, W., Hejnol, A., Vizoso, D.B., Ramm, S.A., Schärer, L., Schatz, M., Hannon, G.J., Ladurner, P., and Berezikov, E. *De novo* genome assemblies of *Macrostomum lignano, Macrostomum hystrix* and *Isodiametra pulchra* shed light on the evolution of pluripotency. *In preparation*.

Simanov, D.*, Mouton, S.*, Arindrarto W., van Nies, K., de Mulder, K., Grudniewska, M., Demircan, T., van der Krogt G., Lengerer, B., Ladurner, P., and Berezikov, E. *De novo* transcriptome assembly of *Macrostomum lignano* and identification of neoblast-specific genes. *In preparation*.

Simanov, D., de Mulder, K., Gommans W., and Berezikov, E. miRNAs of flatworm *Macrostomum lignano*: biogenesis, expression profiling and possible role in stem cell functioning. *In preparation*.

Demircan, T.*, de Mulder, K.*, **Simanov, D.**, Glashauser, J., Vizoso, D.B., Schärer, L., Ladurner, P., and Berezikov, E. Flatworms go transgenic: *Macrostomum lignano* as an emerging stem cell model. *In preparation*.

*These authors contributed equally.