Universidade do Algarve

Molecular mechanisms triggering differentiation of the embryonic endodermal stem cells

Eduarda Mazagão Guerreiro

Dissertação de Mestrado

Mestrado em Ciências Biomédicas

Trabalho orientado por:

- Doutora Natalia Soshnikova, Institute of Molecular Biology, Mainz, Germany

- Professor Doutor José Bragança, Departamento Ciências Biomédicas e Medicina, Universidade do Algarve, Faro, Portugal

2013

Universidade do Algarve

Molecular mechanisms triggering differentiation of the embryonic endodermal stem cells

Eduarda Mazagão Guerreiro

Dissertação de Mestrado

Mestrado em Ciências Biomédicas

Trabalho orientado por:

- Doutora Natalia Soshnikova, Institute of Molecular Biology, Mainz, Germany

- Professor Doutor José Bragança, Departamento Ciências Biomédicas e Medicina, Universidade do Algarve, Faro, Portugal

2013

“Molecular mechanisms triggering differentiation of the embryonic endodermal stem cells”

Mestrado em Ciências Biomédicas

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída

_____

Eduarda Mazagão Guerreiro

Acknowledgements

This project was an opportunity that allowed me to grow professionally in a situation during which the experiences I had, the people I met,

First I have to express my gratitude to my supervisors, Dr. Natalia Soshnikova for giving me the opportunity to develop a project at her group, the group of Epigenetic Regulation of Transcription in Mouse Development and Disease - Institute of Molecular Biology (IMB), Mainz, Germany and guiding me through it; and Dr. José Bragança for having accepted to accompany me (although from far away) in this projet and for being available to help.

As this wouldn't have been possible at all without the support from my parents, I have to express them my gratitude for continue to believe in me and for making this opportunity possible.

When I moved to Mainz to carry out this project Johannes, Jan and Henrik were the first people I met and received me at the house in a great way! Thank you for that guys and for just being there! Also very important during my stay, was the great group of people that received me in the lab – Christina, Stefi, Juri and Elif. They made my arrival very smooth, even though, at some moments, they started to communicate in a very strange language (german)… Until today I am still trying to understand how you guys manage to communicate.

It is also important to me to acknowledge the Institute of Molecular Biology for accepting me there to develop my work. Dr. Andreas Vonderheit, Dr. Malte Paulsen and Ina Schäfer, from IMB core facilities were also very important to the correct development of this project – Thank you.
To the lab neighbors – Thomas, Jonathan, Kat, Thaleia and Stephanos – thank you for the moments spent outside the IMB.

My adopted cousins Ana e Nelson (last, but not the least!), there are no words to thank you for just being there!

i. Abstract

Mouse small intestine is a highly organized tissue with a three dimensional structure with finger-like protrusions – villi – surrounded at the base my multiple invaginations – crypts of Lieberkühn – making it the body largest interface. Due to the constant exposure to harsh conditions, the small intestine has a high cell renewal rate fuelled by stem cells present at the bottom of the crypts. In embryos, the structure that will give rise to the small intestine is formed around e8.5 to e9.0 and progressively acquires the characteristics of the small intestine with villus emerging at e15. At this time, cell proliferation is restricted to the intervillus region and will continue to be located there until the crypts develop in the postnatal period.

This work aimed to 1) validate the transcription profile of the ISCs previously obtained by performing *in-situ* hybridization analysis on paraffin sections of mouse embryos (e12.5, e13.5, e14.5 and e15.5) and adult small intestine and 2) understand in more detail the mechanisms involved in gene regulation by focusing in DNA methylation performing MBD-seq analysis. Bisulfite sequencing protocol was optimized for further DNA methylation analysis.

ISH analysis of Hes1, Notch2, Ascl2, Muc4, Dusp1, Cps1, Kcne3, Igfbp5 and Shh was indicative that the genetic program leading to the adult small intestine change, as target genes were detected to be expressed at different developmental stages. MBD-seq data allowed to detect also changes in DNA methylation in the genes Shh, Grb10, Foxa1, Id2, Ndufaf3, Nrp, Fz2, Meis1, Mdk, Efbn2, Lgr5, Vdr, Olfm4, Elf3, Ephb3 and Oct4. For the majority of the genes, an active transcription was associated with unmethylation of the DNA and methylation increased as the genes were inactive. In respect to the genes where this relation was not possible to establish, transcription regulation as result of other epigenetic markers should be investigated.

Key works: Small intestine, villi, crypt, intestinal stem cells, DNA methylation, epigenetic markers

ii. Resumo

O intestino delgado de ratinho é um tecido extremamente organizado, caracterizado por uma estrutura tridimensional. Esta estrutura apresenta protrusões em direcção ao lúmen – as vilosidades – rodeados na sua base por várias invaginações – as criptas de Lieberkühn. Como consequência desta estrutura, o intestino delgado é a maior superfície do corpo e está exposto constantemente a condições extremamente agressivas. Por este motivo, este tecido apresenta uma taxa de renovação celular extremamente elevada. A renovação do epitélio intestinal é assegurada pelas células estaminais presentes no fundo das criptas. A nível embrionário, a estrutura que irá dar origem ao intestino delgado forma-se no estadio e8.5 a e9.0, a partir da qual adquire progressivamente as características do intestino delgado funcional. As vilosidades surgem cerca de e15.0 e a região de proliferação celular restringe-se ao tecido entre as vilosidades onde permanecem até à formação das criptas após o nascimento.

Este trabalho propôs-se a 1) validar o perfil de transcrição das células estaminais intestinais que havia sido determinado previamente, recorrendo para o efeito a hibridação *in-situ* de secções de embriões de ratinho (e12.5, e13.5, e14.5 e e15.5) e de intestino delgado em parafina, bem como 2) compreender mais detalhadamente os mecanismos envolvidos na regulação da transcrição de genes focando-se na metilação do ADN através de análise por MBD-seq. O protocolo de sequenciação após tratamento com bissulfito de sódio foi optimizado para análises posteriores de metilação do ADN.

A análise por hibridação *in-situ* de slides do genes Hes1, Notch2, Ascl2, Muc4, Dusp1, Cps1, Kcne3, Igfbp5 e Shh produziu resultados que indicam que o programa genético que conduz a um intestino delgado maduro é alterado. Isto é baseado nas alterações observadas na expressão dos genes cujos transcritos eram detectados em diferentes estadios de desenvolvimento. Relativamente aos dados de MBD-seq, estes permitiram detectar alterações na metilação dos genes Shh, Grb10, Foxa1, Id2, Ndufaf3, Nrp, Fz2, Meis1, Mdk, Efbn2, Lgr5, Vdr, Olfm4, Elf3, Ephb3 e Oct4. Para a maioria destes genes foi possível estabelecer uma relação entre o estado de transcrição e os níveis de metilação do ADN que os codifica – nos genes que estavam a ser transcritos a metilação do ADN era reduzida ou inexistente enquanto que os níveis de metilação aumentavam

consoante os genes estavam inactivos. Relativamente aos genes a respeito dos quais não foi possível manter esta relação, seria importante validar a hipótese de que outros marcadores epigenéticos possam estar envolvidos na regulação da expressão.

Palavras-chave: Intestino delgado, vilosidades, criptas, células estaminais do intestino, metilação do ADN, marcadores epigenéticos

iii. List of Abbreviation

A	Adenine
aa	Aminoacid
AE	Adult enterocytes
AP	Alkaline phosphatase
APC	Adenomatous polyposis coli
APRO	Antiproliferative
Ascl2	Achaete-scute complex homolog 2
BCIP	5-brom-4-chlor-3’-indolyl phosphatase
BMP	Bone Morphogenetic Protein
Bp	Base pair
Btg2	B-cell translocation gene 2
C	Cytosine
CBC	Crypt base columnar
ChIP	Chromatin immunoprecipitation
CK1	Casein kinase 1
Co-Smads	Common-mediator Smads
CpG	Cytosine-phosphate-guanine
Cps1	Carbamoyl phosphate synthetase I
cRNA	complementary RNA
DIG	Digoxigenin
Disp	Dispatched
Dll	Delta
DNA	Deoxyribonucleic Acid
DNMTs	DNA methyltransferases
DNP	Dinitrophenol
dNTPs	Deoxyribonucleotide triphosphates
Dvl	Dishevelled
DSL	Delta, Serrate, Lag
Dusp1	Dual specificity phosphatase 1
ECM	Extracellular matrix

E.coli	*Escherichia coli*
EGFP	Enhanced green fluorescent protein
Eph	Erythropoietin-producing hepatocellular
Ephrin	Eph receptor interacting proteins
ERK	Extracellular signal-regulated kinase
ESC	Embryonic stem cell
FACS	Fluorescence activated cell sorting
FBS	Fetal Bovine Serum
Fz	Frizzled
Fw	Forward
G	Guanidine
gDNA	Genomic DNA
GI	Gastrointestinal
Gli	Glioma-associated
GPI	Glycosylphosphatidylinositol
GSK3β	Glycogen synthase kinase 3β
Hh	Hedgehog
IHC	Immunohistochemical
Id2	Inhibitor of DNA binding 2
Ihh	Indian hedgehog
IGF	Insulin-like growth factor
Igfbp5	Insulin growth factor binding protein 5
IPTG	Isopropyl-β-D-thiogalactopyranidose
ISC	Intestinal Stem Cell
ISH	*In-situ* hybridization
Jag	Jagged
JNK	C-jun N-terminal kinase
KAc	Potassium acetate
Kb	Kilobase
kDa	Kilodalton
LB	Luria-Bertani
LEF	Lymphoid enhancer factor
Lgr5	Leucine-rich G protein-coupled receptor 5

LRP	Low-density lipoprotein receptor-related protein
MBD-seq	Methyl-CpG binding domain protein sequencing
MBPs	Methyl-binding proteins
MCS	Multiple cloning site
$MgCl_2$	Magnesium cloride
min	Minutes
mRNA	Messenger RNA
Muc4	Mucin 4
NaOH	Sodium hydroxide
NECD	Notch extra-cellular domain
NEXT	Notch extracellular truncation
NICD	Notch Intracellular domain
Ngn3	Neurogenin 3
nt	Nucleotides
O/N	Overnight
PCR	Polymerase Chain Reaction
pDNA	Plasmid DNA
PFA	Paraformaldehyde
Ptch	Patched
qPCR	Real-time Polymerase Chain Reaction
RNAs	Ribonucleic Acids
R-Smads	Receptor activated Smads
RT	Room temperature
RTK	Receptor tyrosine kinase
Rv	Reverse
SDS	Sodium dodecyl sulphate
Sec	Seconds
Shh	Sonic hedgehog
Smo	Smoothened
Susd4	Sushi containing domain 4
T	Tyrosine
Ta	Temperature of annealing
TA	Transit amplifying

TCF	T-cell factor
UTR	Untranslated region
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Zfp36	Zinc finger protein 36

iv. List of Images

v. List of tables

vi. Table of contents

1. Introduction

Biological systems are in constant search of a homeostatic state. This is the state that offers the best conditions in order for the events needed to sustain life occur in a optimal manner. There are many ways that a system can regulate itself to maintain its homeostatic state while keeping up with the always changing environment. This can include adjustments to pH fluctuations, increase or decrease of secretions for pathogen elimination, just to mention a few. Even the constant renewal of cells in a tissue is important to assure a proper function.

A tissue where cell renewal is a constant is the small intestine. In this structure, the epithelium is the most vigorously self-renewing tissue of adult mammals. The harsh conditions to which the tissue is subjected are the reasons for it[1,2].

1.1 Small Intestine

The small intestine is a highly organized tissue that serve a broad array of functions that are mainly absorptive and protective[3,4]. For the most part of the embryonic development the gut goes through a series of changes. Throughout the different stages, the small intestine accompanies the body's development, continuing until it reaches maturity. While doing so, the gut acquires its final structure the cells lining it start to differentiate and to acquire their final identity[4].

1.1.1 Adult small intestine

The simplest representation of the adult intestinal tract is a tube with a wall composed of three different layers[5] (Fig 1.1 a), each with its own function and characteristics. The

outer layer, formed by several sheets of smooth muscle and by the enteric nervous system ensures the peristaltic movements of the intestine. Connective tissue with numerous blood and lymph vessels, nerve fibers and cells of the immune system makeup the middle layer while the most inner layer, the mucosa, consist of a simple epithelium responsible by digestion and absorption of approximately 95 % of nutrients[2,5,6].

The small intestine tube structure is also complemented with a three dimensional structure characteristic of the lumen. The presence of finger-like protrusions – **villi** – surrounded in the base by multiple invaginations – **crypts of Lieberkühn** – increases the surface area of the gut[2,5-7] (Fig 1.1 b). As a consequence, the small intestine is the biggest surface between the inside and the outside of the body. Although permeable to nutrients and other molecules, it blocks the passage of harmful antigens, bacteria or macromolecules[6,8]. Function of the mucosa epithelium is ensured by a variety of differentiated cells populating it[2,3,9].

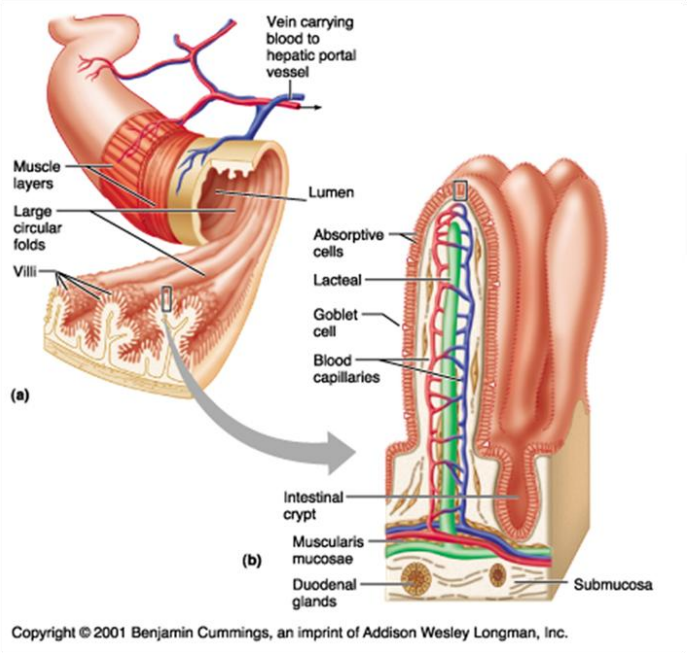


Fig 1. 1: Structure of the adult small intestine in mammals. The small intestine, in its simplest form as a tube with several tissue layers surrounding the lumen (a). In the interior of this structure the villus cover the internal layer as they project towards the lumen as they are surrounded by the crypts of Lieberkühn (b). Image modified from http://www.made4ll.com/anatomy/small-intestine-villi/.

In the villi, three main types of epithelial cells are found: **enterocytes**, **goblet cells** and **enteroendocrine cells**[1] (Fig 1.2). Enterocytes are the most abundant cell type found in the small intestine with their number being higher in the duodenum. The cells are columnar in shape, with a basal nucleus and apical microvilli (that further increase the absorptive surface). Besides being responsible for absorption of nutrients and water by active and passive transport through the epithelium, these cells secrete a cocktail of hydrolytic enzymes into the lumen[2,5,9-11]. Second most abundant cell in the small intestine epithelium are the Goblet (or mucosecreting) cells that are found scattered from the middle of the crypt up to the tip of the villus. They are characterized by having mucigen granules in their cytoplasm, that will be released into the intestinal lumen for lubrication and protection of the mucosa and are found in a higher number in the distal region of the gut[2,3,5,9,10]. The third type of differentiated cell lining the villi are the enteroendocrine. They comprise less than 1% of the total epithelial cells and are distributed evenly along the small intestine. Enteroendocrine cells contain numerous dense granules containing the secreted peptide hormones, that include serotonine, secretine, and substance P, that will control gut physiology[1,5,9,10,12].

The rapid self-renewal of the epithelium is orchestrated by the proliferative compartment, the intestinal crypts. Here, **adult stem cells** and **Transit amplifying** (TA) **cells** will proliferate continuously, generating 250 new epithelial cells per day (Fig 1.2). They migrate upwards from the bottom of the crypts up to the crypt-villus junction, differentiating along the passage. Cells, that now are completely differentiated, continue to migrate upward from the junction until they are exfoliated at the tip of the villus[2,6,8,10,13,14].

The differentiated cells of the epithelium found in the villi, carry out their function of nutrient absorption even though regularly exposed to extremely harsh external conditions. Constant attacks of chemical, mechanic or pathogenic nature requires a high cellular turnover for the villi to maintain its functions. This means a complete epithelium renewal every 3-5 days[7].

A forth type of differentiated epithelial cell, **Paneth cells** (Fig 1.2), are only found in the proliferative compartment. In the bottom of the intestinal crypts, Paneth are columnar cells with of large secretory granules in the cytoplasm. In the granules, specific proteins such as lysozymes, defensins and small molecular weight proteins with

antibacterial function, can be found. All these factors together with the fact that Paneth cells have phagocytic properties are strong indicators that these cells have a role in immunity[1,2,9,10].

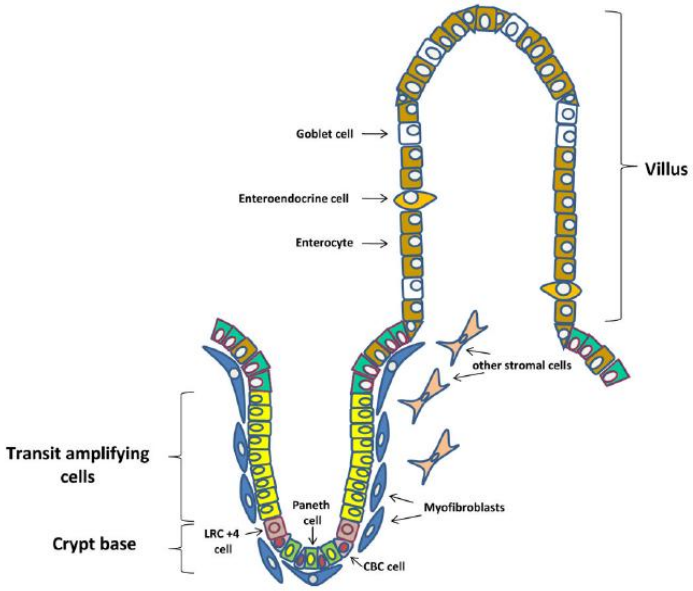


Fig 1. 2: Representation of the crypt-villus axis in the small intestine. This structure can be divided into absorptive and proliferative compartments. In the absorptive compartment, the fully differentiated cells (enterocytes, goblet and enteroendocrine cells) carryout the absorptive functions of the small intestine. In the second compartment, stem cells and transit amplifying cells divide in order to feeding the renewal of the epithelial cells in the villus[15].

In summary, the small intestine epithelium is a very dynamic tissue. The four main cell types are classified into absorptive – enterocytes – or secretory – goblet, enteroendocrine and paneth, according to their characteristics[1]. This lining is a very specialized barrier protecting the organism from external aggressions resulting in the need for a high rate of cell turnover. Fuelling cell renewal are the crypts where 250 cells are produced and differentiate every day in each crypt. To reach this dynamic state, the intestine goes through several developmental stages as it differentiates from the embryonic endoderm[4].

1.1.2 Small intestine in embryos

In the early developmental stages of vertebrate embryos three primary germ layers, endoderm, mesoderm and ectoderm arise after gastrulation[4,16]. Formation of the primitive gut tube starts with endoderm closure in the anterior and posterior intestinal portals continuing toward the middle of the embryo, fusing around embryonic day 8.5 (e8.5) to e9.0[11,17] (Fig 1.3) . Simultaneously with gut tube formation, the epithelium elongates along the anterior-posterior axis and undergoes the patterning of the foregut at the anterior region, midgut and hindgut[4,11] at the posterior end of the gut tube. From these regions, the endoderm derived organs will develop by local swelling, budding or coilling[18]. Foregut will differentiate into thyroid, esophagus, lungs, stomach, pancreas, liver, gall bladder and duodenum while the mid and hindgut give rise to the small and large intestines, respectively[11].

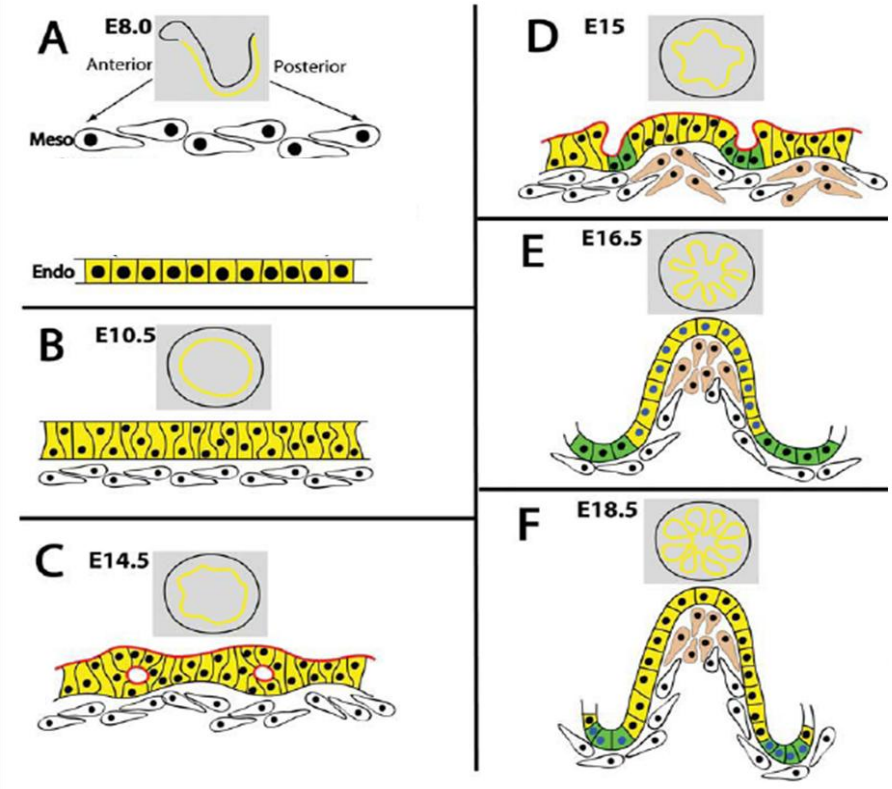


Fig 1. 3: Representation of mouse intestinal development in embryo. From e8.0 to e8.5 the primitive gut tube is formed with endoderm closure (A). Epithelium elongates and forms a pseudo stratified epithelium (B). The primitive gut continues to elongate and at e14.0 the epithelium changes from a thightly packed structure to a columnar organization with intraepithelial cavities starting to form (C). Villus start to emerge at e15 (D) becoming evident at e16.5 (E). From e17, proliferative cells are restricted to the intravillus region – in green (F). Image from Spence *et al* (2011)[4].

Once the gut tube is formed, the simple epithelium condenses into a pseudostratified epithelium at e9.5. Then, and until e13.5, the intestine elongates together with the embryo and both the thickness of the tube walls and the circumference of the lumen increase[4,12]. At e14 the stratified epithelium reorganizes changing from a tightly packed simple epithelium with the nuclei at several levels to a columnar organization with basal nuclei[4]. Simultaneously, secondary lumina (or intraepithelial cavities) start to form in deep basal layers of the stratified epithelium while cells from the mesenchyme start to condensate and this layer begins to invaginate to form the nascent villi[4,12]. Villus emergence is evident in mouse at e15 and progress in a rostral-caudal orientation. At this point cell proliferation occurs along the epithelium, reducing as the embryo develops. Cells proliferate and differentiate into one of the four main types enterocytes, goblet cells, enteroendocrine cell and tuft cells[4]. At e17 proliferating cells are restricted to the intervillus region[12] where they remain until the small intestine continues to develop in the postnatal period when crypts develop[12], hosting the proliferative compartment and the stem cells that fuel it (Fig 1.3).

1.2 Stem cells

The ability to self renewal and to be progenitors of all the body's differentiated cells are features that are used to describe **stem cells**[19,20].

Stem cells are present throughout the different developmental stages of an organism into and during the adult life. **Embryonic stem cells** (ESC) as the name indicate, are found in the embryos and are classified as pluripotent, as they can give rise to all body lineages[21]. **Adult stem cells**, found in adult, fully differentiated tissues, are classified as multipotent. This classification is due to the fact that, adult stem cells can only self-renew and differentiate into the different cells of the specific tissue from which they were isolated[19,21].

1.2.1 Stem cells in adult small intestine

As mentioned in section 1.1.1, the harsh conditions of the intestinal lumen require a high cellular turnover for homeostasis maintenance. There are two key "elements" for the successful cell renewal of the intestinal epithelium: adult stem cells, or **intestinal stem cells** (ISC), and TA cells.The ISC are confined to the bottom of the intestinal crypts and are responsible for continuously providing cells for epithelial renewal. While differentiating into one of the intestinal cell lineages, ISC have to keep their population thus maintaining the proliferative compartment functional[10]. So far, two models have been presented to explain this balance. One of the models, the Deterministic model, suggests that individual stem cells divide asymmetrically; one of the daughter cells is kept at the bottom of the crypt as a stem cell, contributing for ISC population maintenance, while the other migrates into the proliferative zone, becoming a TA cell[10,22]. The Stochastic model, on the other hand, hypothesize that stem cells divide symmetrically and stay at the bottom of the crypt as stem cells; as stem cells continue to divide giving rise to two stem cells they are eventually pushed out of the bottom of the crypt and consequently become TA cells[10,23].

It has been reported that stem cells divide every 24 hours and that there are approximately 10 to 15 stem cells at the bottom of each crypt[22]. With this rates of cell division it would not be possible for the ISC per se to give rise to up 300 cells per day[5]. Once the stem cells leave the bottom of the crypt they pass into the proliferative zone becoming TA cells. The higher division rate these cells present, with cycles every 12 to 16 hours ensure the high number of cells necessary for the epithelium self-renewal and will migrate upwards the crypt walls[5,22]. While doing so, TA cells undergo up to six rounds of cell division and start to commit to one of the cell types, until they exit the crypt fully differentiated [2,5].

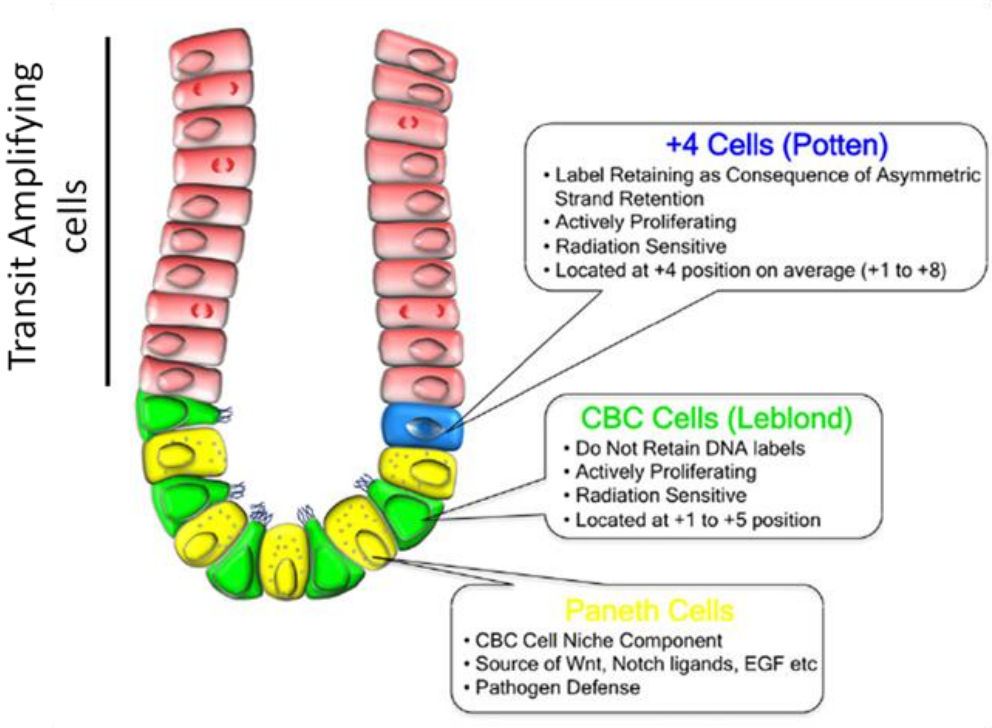


Fig 1. 4: Intestinal stem cells of the small intestine. The specific location of the stems cells is a topic of some discussion with two stem cell populations reported to be present in the bottom of the crypt – crypt base columnar (CBC) cells and the +4 cells. Image from[7].

Although it has been known that the ISC are located in the bottom of the crypt, their specific localization in the crypt has been a topic of some discussion. Two stem cell populations have been reported to be present at the bottom of the crypt, the **crypt base columnar** (CBC) cells and the label-retaining **+4 cells**[24] (Fig 1.4). Behind the description of these two different stem cell populations are the findings of Cheng and Leblond (1974) in respect to the stem cell model and the works from Cairnie *et al* and Potten *et al* (1974) for the +4 model[7].

While performing electron microscopy studies in the crypts of the small intestine, Cheng and Leblond found that other cells beside Paneth populate the bottom of the intestinal crypts[2]. They found slender, immature, cycling cells lodged between the Paneth cells that were named CBC cells[24,25]. Another characteristic of this type of cells is that they are sensitive to tritiated-thymidine exposure. In their studies, Cheng and Leblond also accounted that after exposure the surviving CBC cells phagocyted the surrounding damaged cells. In this process radioactive phagossomes appeared in the surviving cells. They were present initially only in the CBC cells and progressively detected in the differentiated cells as cells divided. Such results were interpreted as evidence of stemness of the CBC cells[7,25,26]. The competing +4 model was initially

proposed when cell tracking experiments predicted that cells in the crypt had a common origin in position 4-5, just above Paneth cells. Additional support to this was given by Potten and colleagues which demonstrated that these cells had desirable stem cells properties. They were found to be sensitive to X and γ-radiation and to have Deoxyribonucleic Acid (DNA) label retention[25]. Sensitivity to radiation would avoid potential carcinogenic genetic abnormalities to be passed throughout the daughter cells. Long-term retention of DNA label resulted from the asymmetric DNA segregation whereby the template, "immortal", strands of DNA were retained by the stem cell in the +4 position, while the newly synthesized DNA strands would be inherited by the daughter cells[2,7,25,26].

One of the reasons for the difficulty on the identification of ISCs is the lack of well validated markers for these cells[3].

1.2.1.1 Stem cell markers

Differentiated cells in the intestinal epithelium have been well defined morphologically and in terms of molecular markers[3]. Unlike these differentiated populations, such definite characterization is not yet possible for the ISCs even though an array of markers have been proposed.

Most efforts were focused in the markers for the +4 population until recently when the stem cell zone resurged. This was possible with the identification of the Leucine-rich G protein-coupled receptor 5 (Lgr5)/GPR49[2]. Since its identification, this gene was included in the list of the potential markers in which are included Bmi 1, mTert, Ascl2, Olfm4, Sox9, Musashi-1[2,26]. Nevertheless, Lgr5 is the most promising with given evidence as a ISC marker[7].

Lgr5 is a 7 transmembrane G-protein with its expression limited to the CBC cells[7]. Cells expressing Lgr5 are able to generate all lineages of the intestinal epithelium *in-vivo*, while *in-vitro* form self-renewing epithelial organoids similar to the crypt-villus axis, as

reviewed by Lin and Barker (2011)[25]. In respect to other possible markers identified so far, complementary studies show that none of them is either expressed exclusively in one of the stem cell populations, or when they are, lineage tracing does not support them[25].

With different cycling dynamics, radiation sensitivity and response to the signaling pathways Lgr5 and +4 cells can be in fact two different populations that complement each other. The CBC cells with a higher cycling dynamic repopulate the crypts, ensuring the daily renewal of the epithelium and the more quiescent +4 cells act as a "reserve" subpopulation that respond in cases of injury or loss of the active stem cells[7].

Nevertheless, with the overlapping expression of Lgr5 with markers associated with +4 population, it has be hypothesized that in fact these two different stem cell populations are the same. In this case the Lgr5+ cells would occasionally acquire +4 characteristics to adapt to changes in the crypt environment[24].

1.2.1.2 Stem cell niche

At the bottom of the intestinal crypt, ISC are enclosed in the stem cells **niche**. This space provides ISC a very specific and dynamic microenvironment that is responsive to environmental signals[15,20].

Beside the ISC, several cell types compose the niche. Epithelial, mesenchymal and smooth muscle cells, basement membrane and macrophages are part of the cellular structure, each contributing in a specific manner[15,27,28]. Myofibroblasts and mesenchymal cells, due to their presence in the adjacent lamina propria are in a location that allows them to regulate cell behavior in the niche through the action of soluble and cell associated factors such as growth factors and cytokines[15,27]. With the combination of this factors ISC are regulated in order to maintain their own population, to proliferate giving rise to the differentiated progeny or even to respond against any perturbations allowing tissue recovery[29].

Small intestine crypts and villi and the diverse cell composition create different regions along the gastrointestinal (GI) tract tightly controlled by a complex interplay of the major signaling pathways[8].

1.2 Cell differentiation/specification

The combination of the environmental factors available for niche and crypt control build a complex signaling network comprising **Wnt/β-catenin**, **Bone Morphogenetic Protein** (BMP), **Hedgehog** (Hh) and **Notch** pathways[27,29], to name a few (Fig 1.5).

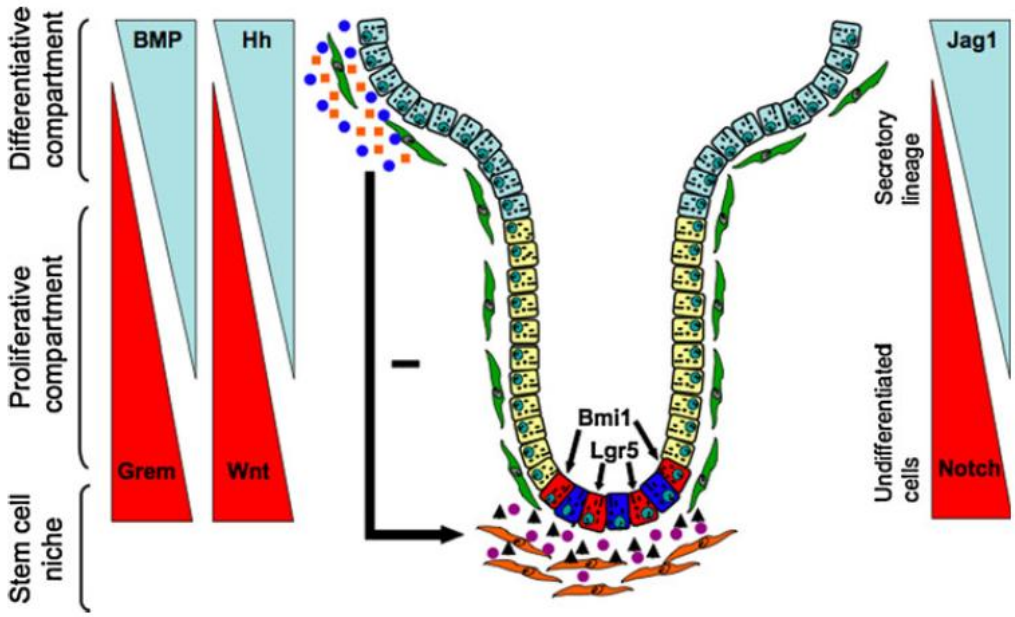


Fig 1. 5: Several secreted factors regulate ISCs and distribution of the TA and differentiated cells. The Wnt/b-catenin pathway is expressed in a reciprocal gradient to the BMP and Hh pathways along the crypt-villus axis, with these last two inhibit the Wnt signaling. Notch expression is higher at the bottom of the crypt and as it gradually decreases towards the villi, ISCs differentiate into the secretory lineage[27].

1.3.1 Wnt/β-catenin pathway

Signaling of Wnt/β-catenin pathway has effects on cell proliferation, cell survival, cell polarity and even cell fate determination[30]. Signaling is ensured by Wnt ligands, a family of secreted glycoproteins evolutionally conserved, found in vertebrates and invertebrates[27,31]. Generally these molecules have approximately 350 aminoacids (aa) with a correspondent molecular weight of ~40KDa and are involved in two major pathways[30,31].

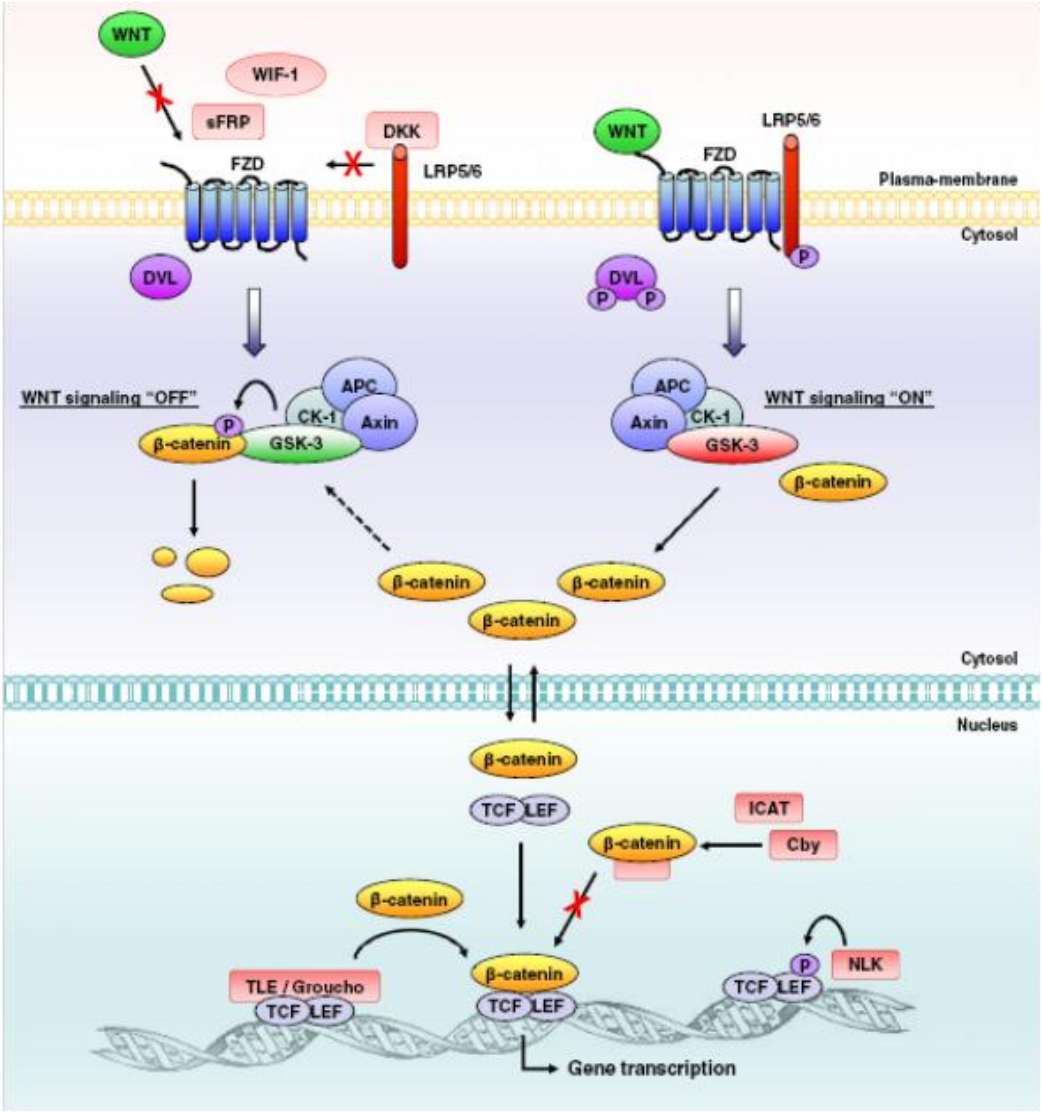


Fig 1. 6: Schematic representation of the canonical Wnt/b-catenin pathway. In the presence of Wnt ligands the ubiquitination of b-catenin is blocked and b-catenin translocates into the nucleus. The expression of Wnt dependet genes is induced[31].

The pathways activated by the Wnt ligand are known as **canonical** (β-catenin dependent) and non-canonical (β-catenin independent)[30]. In the canonical pathway (Fig 1.6), activation occurs when a Wnt ligand binds to the seven-transmembrane **Frizzled** (Fz) receptor and its co-receptor **Low-density lipoprotein receptor-related protein 6** (LRP6) or its close relative LRP5[8,32]. In the absence of the activation signal, β-catenin in the cytoplasm is inactive and it easily captured by the destruction complex formed by Axin, adenomatous polyposis coli (APC), casein kinase 1 (CK1) and glycogen synthase kinase 3β (GSK3β)[8]. CK1 and GSK3β phosphorylate β-catenin in the N-terminal region signaling it for ubiquitination preceded by proteasomal degradation. β–catenin is prevented to reach the nucleus thus repressing Wnt target genes[32,33]. In the presence of Wnt ligands the Wnt-Fz-LRP6 complex formed recruits the intracellular protein Dishevelled (Dvl)[32]. The LRP6 intracellular tail is phosphorilated by Dvl and Axin is recruited to this site inhibiting β-catenin ubiquitination[33]. β-catenin accumulates in the cytoplasm and translocates into the nucleus where it induces the expression of Wnt dependent genes by activating the transcription factors of the T-cell factor/lymphoid enhancer factor 1 (TCF/LEF-1) family[31,33].

In the less studied non-canonical Wnt pathways, Wnt ligand activates the intracellular signaling cascade independently of the LRP5/6[30]. Activation of cJun-N-terminal kinase pathway or Ca^{2+}-dependent pathways occur for the regulation of cell polarity or cell adhesion and motility[30,31].

In the adult small intestine, Wnt/β-catenin is essential for ISC maintenance, proliferation and compartmentalization of the crypt and differentiation zones of the epithelium[4,34]. The Wnt ligands secreted by the myofibroblasts at the basement membrane of basal crypt epithelium creates a signaling gradient of Wnt/β-catenin pathway along the crypt-villus axis, stronger at the bottom of the crypt[5,27]. Target genes of the Wnt/β-catenin pathway include c-myc, involved in cell proliferation, EphB2 and EphB3, controlling segregation that occurs in the crypt, Sox9, a regulator of Paneth cell differentiation, and Lgr5[27], a cellular receptor for the enhancement of Wnt signaling being itself controlled by the Wnt pathway[7,35]. For activation of the Lgr5 amplifying system, a exogenous ligand that make part of a small family of R-spondins, binds to the Lgr5 receptors initiating the intracellular signaling cascade[7,8,35] amplifying the Wnt/β-catenin signaling.

The role of Wnt/β-catenin pathway is to be a master regulator of both embryonic development and adult tissue homeostasis. Nevertheless, studies have shown that Wnt signaling is key for the control of cell proliferation while other pathways regulate other processes in tissue homeostasis. Such case is the Notch signaling cascade who's key role is the suppression of cell fate[36].

1.3.2 Notch pathway

Notch pathway signals for the maintenance of the stem cell signaling network, cellular patterning and cell organization within a tissue[4,37,38]. Notch is a highly conserved signaling network presenting, in mammals, 4 receptors, Notch 1 to 4, and 5 ligands, Delta-like (Dll) 1, 3 and 4, and Jagged (Jag) 1 and 2[36,39,40] .

Both Notch receptors and ligands are cell-surface type I single-pass transmembrane proteins with a extracellular domain[37,41,42]. Activation of the pathway is a short range event through cell-cell interactions between the receptor and the ligand, expressed in the neighboring cells[39,40].

Notch (receptor) is a heterodimer at the plasma membrane formed by the Notch extra-cellular domain (NECD) and a membrane-tethered intracellular domain called NTM[42]. Both domains interact to form the heterodimer through a non-covalent interaction Ca^{2+} dependent[42]. For activation of the signaling cascade, the ligands of the DSL family (Delta, Serrate, Lag) establish a strong ligation with the NECD initiating a process involving proteolysis and endocytosis of the receptor[37]. This ligation cause the complex receptor-ligand to be available for cleavage by extracellular ADAM proteases[36,42]. The complex is sliced and the ectodomain of Notch is released while a membrane-bound form of Notch, called Notch extracellular truncation (NEXT) is activated. NEXT is further processed through the action of the intramembranous protein complex γ-secretase[36,42]. This complex cleaves NEXT at two endomembrane sites resulting in the release of Notch intracellular domain (NICD) to the cytoplasm[42]. Once in the cytoplasm, NCID translocates to the nucleus were it interacts with the DNA

binding protein CSL family and recruits the coactivator Mastermind initiating the transcription of Notch downstream target genes[37,41] (Fig 1.7).

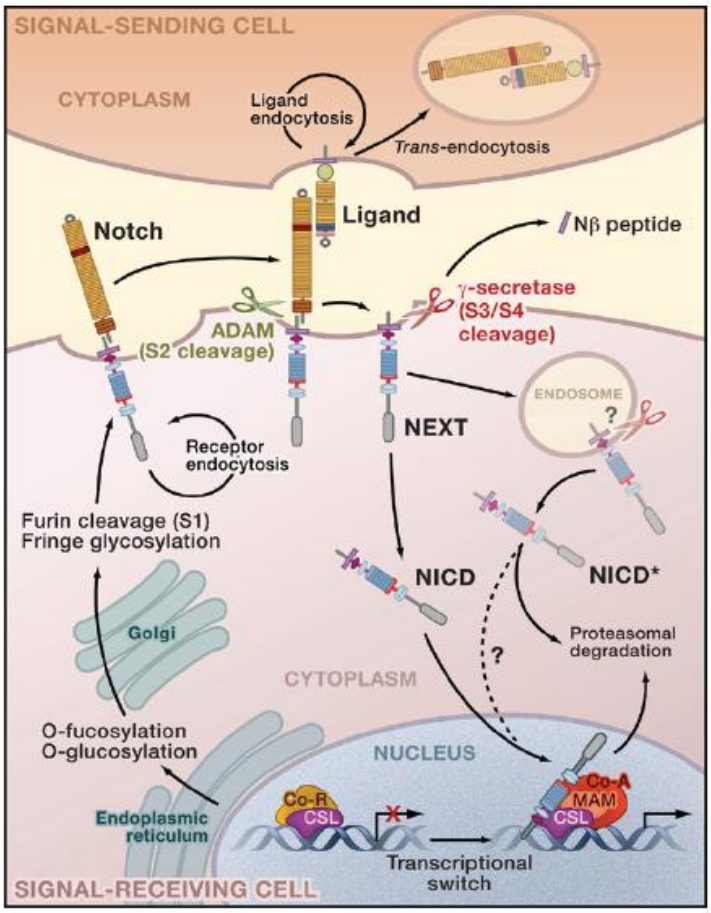


Fig 1. 7: Model for the Nocth signaling pathway. The binding of the DSL family ligands initiate the signaling cascade triggering the transcription of Notch downstream target genes. Image from [41].

Notch downstream target genes often differ among cell types. Nevertheless, the most widely expressed and best characterized target genes of Notch signaling cascade belong to the HES/HEY family. One of the major functions of these proteins is the repression of transcription factors of the Ascheate-scute like 1 and Neurogenin family. Notch has been shown to regulate expression of genes like myc, cyclinD, CDK5 in cell proliferation, or cell cycle inhibitors like p21 in cell types that Notch promotes differentiation[39].

In the small intestine of mammalian, due to the characteristics of this tissue, Notch signaling pathway has a dual role[39]. It controls the differentiation of daughter cells into enterocytes of the absorptive lineage while, together with Wnt pathway, Notch promotes stem cell proliferation[3,39,40]. In the adult small intestine Notch components are expressed in the epithelium at the bottom of the crypt in the region of the stem cell[3,40].

On the other hand, in the developing intestine, at E13.5 and E18.5 of the mouse, Notch ligands Jag1, Jag2 and Dll1 are predominantly expressed at the mesenchyme[40].

1.3.3 BMP pathway

BMPs are a family of conserved growth factors and form the largest subgroup of the Transcription Growth Factor Beta (TGFβ) superfamily[43,44]. They are synthesized as large 400-500 aa precursors and are cleaved into 50-100 aa active proteins with seven cysteins, of which 6 will form 3 intramolecular disulfide bonds[43,45]. The remaining cysteine is key for dimerization with another monomer, forming a covalent disulfite bond between both molecules producing homo or heterodimers[43].

The role of BMPs is very wide. During embryonic development they influence gastrulation, neurogenesis, apoptosis and hematopoiesis[45]. In adult, these proteins are involved in homeostasis maintenance participating in tissue remodeling and regeneration regulating stem cell properties[46].

The BMP canonical signaling pathway functions through a receptor mediated intracellular signaling[46]. The signaling cascade is initiated when the extracellular BMP ligands bind to the receptor[47]. BMP receptors are formed by type I and type II receptors in such way that different combinations of the type II with one of the three type I receptors may determine the receptors specificity[46]. When BMPs bind to the receptors, it is presumed that a conformational alteration of the type II receptor occurs thus activating, by cross-phosphorylation, type I receptor[43-45]. In turn, the active type I receptor transiently associate with receptor-activated Smads (R-Smads, Smad1/5/8) phosphorylating and activating it[45,47]. R-Smads are quickly released from the receptor and interact with the common-mediator Smads (Co-Smads, Smad4). Two R-Smads and one Co-Smad form a complex that translocates from the cytoplasm into the nucleus modulating the transcription of target genes[44,46].

Although not as well studied as the Smad-dependent pathway, BMPs are able to initiate other pathways such as extracellular signal-regulated kinase (ERK), map kinase p38,

C-jun N-terminal kinase (JNK), and nuclear factor kappa beta (NFκB). Through these pathways, BMPs are involved in cell survival, apoptosis, migration and differentiation[43].

In the small intestine several pathways are involved in the regulation of stem cell self-renewal with each cascade playing either different or overlapping roles[46]. In the case of the BMP pathway, this is involved the regulation of stem cell proliferation by inhibiting it[27]. The BMP ligands 2 and 4 mainly synthesized and released by the cells of the intervillus mesenchyme at the villus tips. They decrease in number towards the crypt and BMP receptors found along the villus epithelium[13,27] BMP signaling overlaps with Hh and Wnt pathways[13,27]. By mediating the action of the Hh pathway formation of ectopic crypts is avoided, thus restricting crypt numbers, while, by suppressing Wnt cascade signaling stem cell, self-renewal is balanced[13,27].

1.3.4 Hh pathway

The Hh proteins provide a critical signal for the proper embryonic development to occur[48]. During these critical developmental stages, Hh proteins are involved in the patterning of the neural tube, lung, skin, axial skeleton as well as the GI tract[48]. In the adult tissues, on the other hand, Hh are still present even though only involved in stem cell maintenance, proliferation and differentiation[48,49].

In mammals there are 3 members of the Hh family: Sonic Hedgehog (Shh), Indian Hedgehog (Ihh) and Desert Hedgehog (Dhh)[50]. They are synthesized as precursor proteins of about 45 kilodalton (kDa) that undergo autocleavage yielding an amino-terminal polypeptide and a carboxy-terminal polypeptide[51,52]. The N-terminal polypeptide, now with about 19 kDa, is modified with a cholesterol molecule added in the C-terminus and a molecule of palmitic acid at the N-terminus[48,52]. With this dual lipid modification, Hh proteins have their membrane association enhanced also affecting their secretion and action range[52]. This culminates with the secretion of the Hh proteins from the cell through the activity of the multipass transmembrane protein Dispatched (Disp)[52].

Until the Hh ligands binds to Patched (Ptch), a 12-span transmembrane protein receptor, the Hh cytoplasmatic signaling cascade is inhibited by the repression of another transmembrane protein, Smoothened (Smo)[50,53]. When a Hh ligand associates with the receptor Ptch, this complex is internalized resulting in the loss of activity of Ptch and Smo repression ends, initiating the signaling[49,51].

The de-repressed Smo signals for the activation of STK36 serine/threonine kinase[53]. In turn, the formation of the GLI degradation complex is inhibited, preventing the phosphorylation of the glioma-associated (Gli) family proteins (Gli 1 to 3) by the complex and consequent ubiquitination and degradation[50,53]. Proteins of the Gli family are zinc finger transcription factors that when are not degraded translocate to the nucleus regulating the transcription of genes such as Gli1, Ptch1, CCND2, FOXL1, and JAG2[50,53].

In the small intestine, the relevance of Hh pathway starts very early, being critical for the normal development and for the establishment of villus and crypts. During the development of the mouse embryo, the epithelium expresses Shh and Ihh, signaling for the receptors in the mesenchyme[13,54]. As the intestine develops, the expression of these ligands becomes restricted to the intervillus regions[13]. In the adult small intestine, expression of Shh and Ihh is maintained in the regions that derive from the intervillus regions – the crypts, where proliferation is present[13]. Its presence here is of great importance for the inhibition of the Wnt canonical pathway, thus regulating stem cell proliferation in the crypts[52,53]. Also, Hh pathway signals for the regulation of the BMP pathway controlling the crypt population[52].

1.3.5 Eph/ephrin pathway

Erythropoietin-producing hepatocellular (Eph) receptors are type-I transmembrane protein and, together with its ligands, comprise the largest subfamily of receptor tyrosine kinase (RTK) with 14 members[55-57].

Relevance of Eph receptors in homeostasis maintenance can be noted by the biological activities they have been implicated so far. Eph receptors are reported to have effects on

actin cytoskeleton, cell-substrate adhesion, cell shape and cell movement, as well as cell proliferation, survival, differentiation and secretion[58,59].

The extracellular domain of the receptor (N-terminal) contain a high affinity binding site for interaction with the ligands, Eph receptor interacting proteins (ephrins)[58]. Ephrins are, in turn, also cell surface-bound proteins that, due to structural differences are divided in two classes: class A, anchored to the plasma membrane through a glycosylphosphatidylinositol (GPI); and class B, are transmembrane proteins with a small cytoplasmatic tail[56,60]. When the ephrins bind to the respective receptor, phosphorylation of the kinase domains occur, at the intracellular region[57,60]. Conformation of the receptors and ligands occur as well as multimerization and clustering and the active kinase phosphorylates other molecules triggering the downstream cascade[57,60].

One of the major distinctive characteristic of the Eph/ephrin signaling is the generation of bidirectional signaling: both receptors and ligands are able to transduce a signaling cascade activated upon receptor-ligand biding[56,58]. This property, together with the multimerization of the Eph/ephrin complex result in its biological roles[60]. As the Eph receptors phosphorylate downstream molecules, in turn signal through a wide set of pathways. The cascades that are activated include H-Ras and MAP kinase pathway (transcription regulation, proliferation and cell migration), Jak/Sak pathway and PI3K pathway (promoting proliferation)[60]. This highlights how the Eph/ephrin is a part of a very complex network of signaling cascades[56].

In the small intestine, epithelial homeostasis is mainly regulated by the Wnt pathway, in part by controlling the balance in the expression of Eph receptors and ligands. As the TA cells migrate, they move away from the Wnt signaling, consequently, EphB expression decreases and ephrin-B expression increases[58].

As described in the previous sections, signaling pathways are key for the maintenance of tissue homeostasis, ensuring that stem cells are maintained, proliferate and differentiate in a orderly manner. It is important to note that some of these pathways begin to be expressed at early developmental stages. In doing so, they contribute to the correct development of the animal.

The control of gene expression during the several developmental stages has to be tight to ensure a balance between stability and plasticity[61]. To this effect DNA post-transcriptional modifications contribute to gene regulation[61]. These modifications are termed **Epigenetic modifications**.

1.4 Epigenetics

As cells multiply and differentiate, either during embryonic development or adult stem cells, gene expression modifies. Some of the genes are activated while others are deactivated[62]. This results in a tight gene expression program specific to a population of cells at a given developmental stage allowing a balance between stability and plasticity[61]. Nevertheless this all happens without changes in the DNA sequence. The reprogramming that is carried out is a result of **epigenetic modifications** to the DNA[62].

Epigenetics has been defined as changes in gene expression that do not result from changes in nucleotide sequences and are mitotically and/or meiotically heritable[63-65]. They are under the effects of environmental cues and are key for regulation of tissue and cell-type specific gene patterns[63,66] (Fig 1.8).

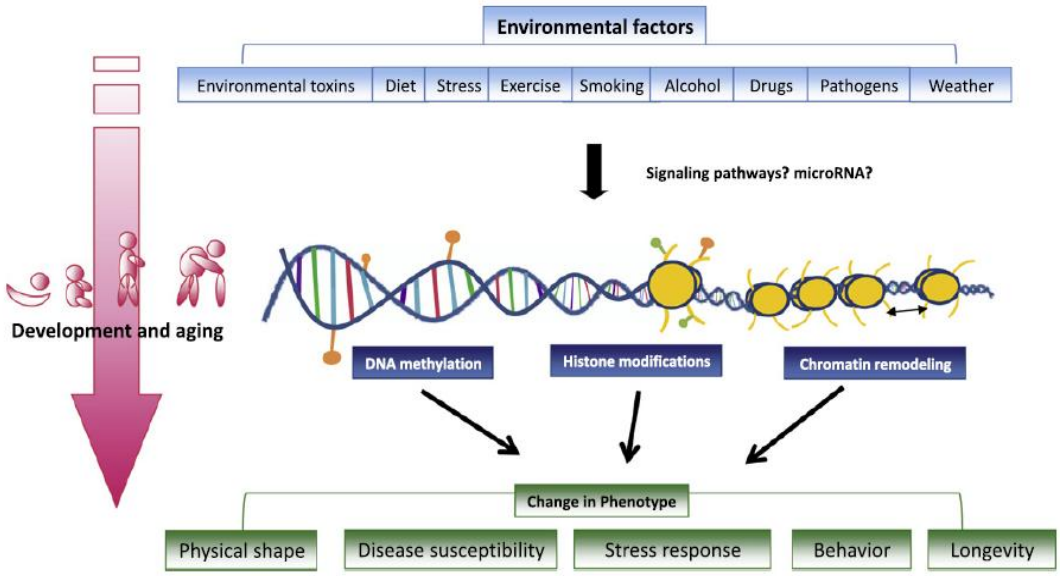


Fig 1. 8: Epigenetic mechanisms and the link between environmental factor and phenotipical alterations[66].

1.4.1 Epigenetic Modifications

One of the challenges in understanding epigenetics is the balance required, each time a SC divides, how the factors will define whether a daughter cell differentiate or maintains its plutipotency[65].

This process is triggered at the very early stages of an organisms life. Shortly after fertilization in mammals both (highly specialized) parental genomes combine to give rise a totipotent embryo, through epigenetic rearrangements[62].

The best-known epigenetic modifications are DNA methylation, changes in chromatin due to histone modifications and the less understood microRNA[63-66].

1.4.1.1 DNA methylation

The best known epigenetic modification is **DNA methylation**[65]. It's the result from the covalent binding of a methyl (CH_3) group to a cytosine in the 5' position, at a cytosine in the dinucleotide cytosine-phosphate-guanine (CpG)[67,68].

Changes in DNA methylation pattern occur through life span. It starts upon fertilization, when gametes methylation signatures are erased and *de novo* patterns are established[66]. At this point, DNA methylation is very low in the stem cells, increasing throughout cellular differentiation[61].

In the majority of mammal cells, 70 to 80% of the CpGs found in the genome are methylated, despite the CpGs that are isolated, clustered or found in a CpG island (DNA region with a high concentration of the dinucleotides)[65,67,68].

Nevertheless, the main indications that DNA methylation is involved in the control of gene expression, come from the fact that certain regions are protected from methylation. The majority of those regions are CpG islands in the 5' end of 60% of human genome that usually comprise promoter and first exon of a gene[65]. When one of this regions

become methylated during the developmental or differentiation stages, long term gene expression is silencing occurs[65,67].

The reaction leading to cytosine methylation is catalyzed by DNA methyltransferases (DNMTs). These enzymes are responsible for maintenance, DNMT1 – ensuring the maintenance of methylation patterns through the cell division; or *de novo* methylation, DNMT3A and DNMT3B, that are able to set new methylation marks in unmodified and hemimethylated cytosines[62,66,68].The presence of the methyl groups in the DNA sequence is believed to prevent the transcription factors to bind to it or to facilitate the assembly of repressor complexes[69].

Detection methods for the study of methylation patterns in gDNA include bisulfate sequencing, methylation-specific PCR, restriction landmark genomic scanning and differential methylation hybridization[65].

Although important, DNA methylation is not the only factor in the repression of gene transcription. **Histone modifications** is also needed for this regulation mechanism[62].

1.4.1.2 Histone modifications

In order to keep the long DNA strands packed in the cell nucleus, associated proteins are required. This complex system of DNA and proteins make up chromatin and it has the ability to alter transcriptional activity as result of modifications on its structure[65]. The basic structure of chromatin is the nucleosome[65]. Each nucleosome is a octamer with two of each of the four core histones, the proteic building blocks of DNA packaging: H2A, H2B, H3 and H4[66,68]. These proteins have two domains, global and tail. The latter is the main target of posttranslational modifications[63]. Such modifications are the key for chromatin condensation, replication, DNA repair and transcriptional regulation and are the result of changes in the ionic charge of the histones tail or by serving as platform for other proteins to bind into[63,66]. Among the histone modifications that include acetylation, methylation, phosphorylation, ubiquitination and sumoylation, the most well studied in the field of Epigenetics, are the changes in methylation and acetylation of histones H3 and H4[66,67].

Much like with DNA methylation, chromatin structure changes throughout development and/or differentiation. It starts as a permissive structure in multipotent cells that progressively closes during differentiation[69].

The study of histone modification, the chromatin immunoprecipitation (ChIP) method is key. In this method the cells are treated with formaldehyde to crosslink DNA and histones. This ligation will be removed after gDNA is fragmented and the fragments that are not bound to DNA-binding proteins are removed from the solution. Two downstream methods were developed to identify the potential sequences associated to the histones[65]. One of the methods, known as ChIP-chip uses microarrays to identify, through hybridization, the DNA regions which were bound to histones. The other method uses the new sequencing technologies to directly identify the DNA regions that were associated to the histones[65].

1.4.1.3 MicroRNAs

MicroRNAs are endogenous noncoding RNAs of 19 to 25 nucleotides (nt) long[65]. These small single stranded molecules, although not coding for proteins, regulate mRNA translation into proteins. This regulation occurs when one of the microRNAs bind to the 3'-UTR (untranslated region) of a given mRNA, preventing the protein synthesis[68]. Although microRNAs act downstream from the other epigenetic factors (DNA methylation, histone modifications) they are epigenetic factors and are under the influence of certain environmental factors[66,68]. Studies indicate that there is a balance established between microRNAs, DNA methylation and histones. For example DNA methylation and histone modifications are able to regulate microRNA's transcription, while some microRNAs can regulate the expression of DNMT3A and DNMT3B as well as Polycomb group genes (key enzymes for regulation of histone modifications)[66].

2. Objective

The study that this thesis reports to, consists in one step of a project on the genetic and epigenetic changes in the small intestine in mice. The project has been and continues to be developed at the Group of Epigenetic Regulation of Transcription in Mouse Development and Disease - Institute of Molecular Biology (IMB), Mainz, Germany.

2.1 Project background

As the small intestine develops and matures, several gene cascades are activated, deactivated or gradually restricted to a small niche. This project's broader goal is to understand gene expression in the stem cells during the developmental process.

The first approach for this project consisted in the isolation of intestinal stem cells from embryo small intestine, e12.5, and adult small intestine, specifically Lgr5 producing cells. Once the cells were isolated, RNA was extracted and sequenced, thus providing important information about genes that were being expressed at those moments.

Based in RNA-seq information, validation of the data was necessary, which was developed by C. Krienke. In her study, some of the genes differentially expressed, according to the RNA-seq data, were selected and *in-situ* hybridization and real-time Polymerase Chain Reaction (qPCR) were performed, successfully validating RNA-seq data.

The study by C. Krinke was a corroboration of the main purpose of this project, allowing the pursue of answers to broader questions such as:

- How does gene expression change during the embryonic development of the small intestine?

  In a embryo, the small intestine, starts as a very simple structure that as the embryo develops acquires a three-dimensional arrangement, accompanied with cell specification. It will be of great interest to understand how these processes modify throughout the developmental stages until adulthood.

- Which are the genetic mechanisms that control the activation and deactivation of the gene networks involved in the development and differentiation?
  Regulation of gene expression is dependent of several markers that, although not altering the DNA structure, alter the availability of certain regions of the molecule to the binding of transcription factors. Taking this into consideration in the developmental stages, having new data regarding this processes could be valuable to understand how stem cells maintain their characteristics and differentiate into mature cells.

2.2 Goal

In order for an organism to reach a mature and fully functional state, a series of events had to occur. The transition from a fertilized egg into the complex organism in the adult stage, is the result of a continuous process controlled by key factors in a genomic level. As new research approaches are being followed, they contribute to understand that the genomic information is dependent of other factors for its regulation – Epigenetic markers.

The goal of the present study was the analysis of genomic DNA (gDNA) methylation in the stem cells of the small intestine in different developmental stages: embryonic and adult. By focusing in the embryonic stages e12.5 and e14.5, and in the adult ISC and fully differentiated enterocytes we aimed to detect the changes in the methylation pattern and attempt to find an association to the changes in RNA expression previously detected.

3. Methodology

This study focus on the *in vivo* differences of stem cells found in the small intestine of mouse embryo and adult. In order to gather more information about the processes involved in stem cell maintenance and differentiation, cells were isolated from untreated mice.

In this experience adult male of the Lgr5-EGFP-ires-creERT2 (Jackson Laboratories) strain were the source of the adult ISC. The choice of this strain was based in the Lgr5-EGFP construct in which the expression of EGFP is under the control of Lgr5 promoter, allowing an easier identification of the adult ISCs. On the other hand, for the collection of endodermal embryonic stem cells, C57BL/6 females were sacrificed and the embryos collected and dissected at the proper developmental stage (established by vaginal plug).

Histological sections were prepared to assess morphological development and correlate the developmental stage with the activation or deactivation of gene expression. *In-situ* hybridization (ISH) was performed at e12.5, e13.5, e14.5 and e15.5 and also in sections of adult small intestine.

To analyze changes in the gDNA methylation, stem cells were isolated from adult and embryo (e12.5) small intestine, through fluorescence activated cell sorting (FACS). From the isolated cells, gDNA was isolated and purified for further analysis by Methyl-CpG binding domain protein sequencing (MBD-seq) and Bissulfide conversion.

3.1 Sampling

Mice, male or female, were sacrificed by hypoxia with CO_2 followed by cervical dislocation.

Male Lgr5-EGFP-ires-creERT2 were dissected. Dissection was initiated with a small abdominal incision that was gradually increased until whole abdominal cavity exposure.

Small intestine was isolated from the GI tract and placed in a petri dish with PBS[1] 1X. The small intestine was cut longitudinally, washed in PBS 1X several times to remove maximum of fecal matter. Tissue was sectioned into smaller fragments and prepared for further processing – ISC isolation or histology.

Following euthanization, female mice were also dissected in order to expose the abdominal cavity. Uterus was removed and placed in PBS 1X. Embryos were dissected under a Leica M80 (Leica Microscopie + Systeme GmbH, Germany). Embryos, in the amniotic sac, were exposed, collected and placed in PBS 1X. Embryos were decapitated and further processed either for embryonic ISC isolation or histology.

3.2 Histochemistry

In this study, the histochemistry technique used was the *in-situ* hybridization. This approach is used as tool to study gene expression in tissues and cell types[70]. Another of its valuable features lies in the spatial information retrieved in respect to gene expression. In histological sections the cells with the target messenger RNA (mRNA) will be identifyed[71].

The process through which ISH detects the presence of a target mRNA in a sample is by the annealing of a complementary RNA (cRNA) to the target in the cell expressing it. As the cRNA chains are labeled with a detectable tag, when they became visible, it is possible to assess the distribution of the target mRNA[71]. Initially tags were radioactive labels and detection was only possible through autoradiography[71]. New non-radioactive labels have been developed with the advantage of being non-radioactive. These labels use chemical modifications to the nucleotides used in the synthesis of the complementary chain. Such chemical modifications include digoxigenin (DIG), dinitrophenol (DNP) and fluorescein[71]. For detection, probes labeled with DIG have to be processed using immunohistochemical (IHC) techniques, with anti-DIG antibodies coupled to an enzyme that is able to catalyze a cromogenic reaction when a subtract is

[1] *Vide* Annex for PBS preparation
[2] *Vide* anex for plasmid vector map and reference points

added allowing the detection of the target mRNA[71,72]. On the other hand, probes that use fluorescein as marker, have the advantage of direct visualization[72].

In this study, ISH was performed in paraffin embedded sections of mouse embryos (at different developmental stages) and adult small intestine. Probes were prepared using DIG as a marker and detection was possible using anti-DIG antibodies coupled with alkaline phosphatase.

3.2.1 Tissue processing

Embryos and adult small intestine collected as described in section 3.1 were initially fixed, in order to maintain tissue structure while ensuring nucleic acids stability[72]. This was followed by tissue dehydration and paraffin embedding previous to paraffin block preparation.

In order to the fix the tissues, samples were placed in 4% paraformaldehyde (PFA) (Sigma Aldrich, Germany) at 4ºC overnight, followed by two washes in PBS 1X, 10 minutes (min) each. Dehydration was performed with 2 incubations, 2 hours each, in 70% methanol (Carl Roth GmbH, Germany) in PBS 1X followed by 2 incubations 100% methanol (Carl Roth GmbH, Germany) also two hours each. Paraffin embedding was initiated by transferring samples to xylol (Carl Roth GmbH, Germany) for 15 min, twice, a 2 hours incubation in paraffin (Carl Roth GmbH, Germany) at 58ºC, followed by an overnight incubation at 58ºC in fresh paraffin. Once the samples were embedded in paraffin, blocks were prepared in a Leica EG1150 H Paraffin Embedding Center (Leica Microskopie + System GmbH, Germany) with special attention to embryo positioning, to allow sectioning to be preformed along an anterior to posterior orientation.

Serial sections, 5 to 6 from each stage, 10 μm thick, were obtained from each block in a Leica RM 2255 Microtome (Leica Microskopie + System GmbH, Germany) and mounted in SuperFrost Ultra Plus (Thermo Scientific, Germany) glass slides. As sections were being prepared they were observed under a Leica DM750 microscope. This ensured that the sections obtained were restricted to the embryo's abdominal

region containing the small intestine. Slides were allowed to dry at room temperature (RT) and stored, also at RT, until required.

3.2.2 *In-situ* Hybridization

The *in-situ* hybridization technique used a complementary RNA strand labeled with digoxigenin, that hybridized with its complementary mRNA sequence present in the tissue. The strategy used for this technique was the use of an anti-digoxigenin antibody that bind to the probe's label and is conjugated with an alkaline phosphatase (AP). Addition of BCIP (5-brom-4-chlor-3'-indolyl phosphatase) allowed AP to produce an intermediate that underwent a dimerization producing an indigo dye with the release of H^+. The free ions reduce NBT (nitro-blue tetrazolium chloride) to NBT-Formazan thus indicating the transcript location in the tissue.

Preparation to the ISH on slides included fragment amplification of the target gene, fragment ligation to a cloning vector - plasmid, transformation of competent cells with the ligation product and culture, small scale plasmid DNA (pDNA) preparation (Miniprep), sequencing, large scale pDNA preparation (Midiprep), pDNA digestion, reverse transcription and probe labeling and probe purification.

Genes chosen to be analyzed in this study were Mucin 4 (Muc4), Sushi containing domain 4 (Susd4), Insulin growth factor binding protein 5 (Igfbp5), Dual specificity phosphatase 1 (Dusp1), B-cell translocation gene 2 (Btg2), Achaete-scute complex homolog 2 (Ascl2), Inhibitor of DNA binding 2 (Id2), Rik, Notch2, Zinc finger protein 36 (Zfp36), Shh and Neurogenin 3 (Ngn3).

3.2.2.1 Fragment amplification

From the RNA-seq tracks obtained as result of the first tasks of this project, target genes were selected based in the expression profiles. Genes with higher expression in one of the stages, e12.5 or adult ISC, were selected to be analyzed by ISH.

With the exception of the Ngn3, all the gene fragments had been amplified and cloned by elements of the laboratory. For this reason, amplification protocol will only be described focusing on the Ngn3 gene.

In order to amplify a fragment of the gene Ngn3, it was required to used the Polymerase chain reaction (PCR) technique. PCR refers to an *in vitro* process that enables the generation of large amounts of a specific DNA sequence, from a small amount of template DNA. For this reaction to occur its required the presence of two synthetic oligonucleotide **primers** that flank the target sequence and are complementary to the opposite strands, with the 3' hydroxyl ends oriented toward each other; a **DNA template** containing the target sequence to be amplified, in this study, gDNA; a **thermostable DNA polymerase** able to withstand multiple cycles of heating up to 95ºC; the four deoxyribonucleotide triphosphates, **dNTPs**, (Adenine, A, Tyrosine, T, Guanidine, G, and Cytosine, C).

A initial **denaturation** step starts the process, with the temperature being raised to about 95ºC to denaturate the template DNA eliminating secondary structures. Synthesis cycles are also initiated with a denaturation step that is followed by a **renaturation** or **annealing** step. In this second step the reaction temperature is slowly cooled to temperatures around 55ºC (varying according to primer optimization) allowing for the primers to pair in the template DNA chain. When the ligations are stable, the **synthesis** or **extension** step is initialized. In order to the synthesis of the new DNA fragments to take place, reaction temperature is raised to ≈72ºC, maximizing catalytic activity of *Taq* DNA polymerase to extend the chains[73]. The repetition of these steps in a limited number of cycles (usually 20-35) will increase the amount of target DNA in a exponential manner, as the amount of DNA duplicates in each cycle.

3.2.2.1.1 Primer design

Design of the specific primers for the Ngn3 gene was performed using the software Primer Premier™ (Biosoft International). Several potential primers were produced. The pair of primers that was chosen was the one with less or most instable secondary structures (such as dimmers and hairpins), low probability of false priming and targeting other regions on the DNA, which was determined by sequencing analysis.

The sequences of the primers used for the Ngn3 are presented in **table 3.1**.

Table 3. 1: Sequences of the specific primers, forward (Fw) and reverse (Rv) for amplification of Ngn3 fragment by PCR and predicted size of the fragment (in base pairs, bp).

Primer	Sequence (5'→ 3')	Size (bp)
Ngn3_Fw	CACTCAGCAAACAGCGAAGA	608
Ngn3_Rv	AAAGCCAGAAGAGGCGAGAT	

3.2.2.1.2 Amplification conditions

The cycling conditions for the PCR amplification of the Ngn3 fragment were as described in **table 3.2**.

Table 3. 2: PCR cycling conditions for amplification of a Ngn3 fragment with specific primers. Data refers to temperatures (in ºC), duration of each step (in min or seconds, sec) and also the number of cycles to each set of denaturation, annealing and extension steps.

Temperature (ºC)	Duration	Number of cycles
94	4 min	
94	1 min	
65	1 min	x 2 cycles
72	1min	
94	30 sec	
65	30 sec	x 33 cycles
72	30 sec	

Reaction mix for Ngn3 PCR was as indicated in **table 3.3**.

Table 3. 3: Reaction Mix for Ngn3 fragment amplification by PCR. Primers were those described in table 3.1 (Sigma-Aldrich, Germany), *Taq* polymerase and Reaction buffer were from the EpiMark Hot Start *Taq* kit (New England Biolabs, Germany) and template gDNA from adult mouse tails digested with proteinase K.

Reaction components	Volume per reaction (µl)
EpiMark Hot Start *Taq* Reaction Buffer 5X	5
dNTP Mix (10 mM)	0,5
Primer Fw (10 µM)	0,5
Primer Rv (10 µM)	0,5
Taq DNA polymerase	0,25
gDNA	0,5
miliQ water	17,75

PCR reaction was carried out as described previously in a thermocycler (Thermocycler TProfessional TRIO, Biometra GmbH, Germany). Products of the PCR reaction were analysed by eletrophoresis in a 1% agarose (AppliChem GmbH, Germany) gel stainned with ethidium bromide (AppliChem GmbH, Germany). To the PCR reaction tubes 10 µl of 6X Orange DNA Loading Dye (Fermentas) were added, and the final volume was loadded into the gel. Amplicon visualization was carried out in a Chemidoc XRS system (Biorad, Germany).

The electrophoresis while confirming the success of the reaction in respect to the fragment of interest, also allowed the separation of the target DNA from primers, dNTPs, and/or non-specific products.

3.2.2.2 Fragment purification

The DNA fragments that, in the electrophoresis gel, had the expected size were excised from the gel and purified using PeqGOLD Gel Extraction kit (PeqLab, Germany). Purification was carried out according to the manufacturer's instructions. Briefly: the excised fragments were placed in centrifuge tubes and weight. To each 0.1g of gel, 100 µl of binding buffer and the mix was incubated (55 to 65 ºC) until the gel was completely melted. A column, provided in the kit, was placed in a 2 ml collection tube and a maximum volume of 750 µl of gel and buffer were transferred and loaded into the column. The column and the tube were centrifuged for 1 min at 10000g and the flow-through discarded. This step was repeated until all the gel and buffer mixture was loaded into the column. Then, 750 µl were added to the column and centrifuged for 1 min at 10000g and the flow-through discarded. This step was performed twice. The columns were dried with a centrifugation of 1 min at 10000g. Columns were transferred to a fresh tube and 50 µl of elution buffer were loaded into the column. Samples were collected with a centrifugation for 1 min at 5000g.

Following purification, samples were quantified using the spectrophotometer Nanodrop (Thermo Scientific, Germany) prior to ligation to the cloning vector.

3.2.2.3 Cloning

Molecular cloning is the process during with a fragment of DNA is inserted in a cloning vector, such as a plasmid, in order to increase the number of copies of the fragment of interest. This is usualy achieved by inserting the vector is a host. The most commonly used are bacteria[74].

A plasmid cloning vector possesses speficic features that include a small size, needed to maximize the transfer efficiency of exogenous DNA into the host bacteria like *Escherichia coli* (*E.coli*), as it decreases with plasmids that are larger than 15 kilobase (kb) long; a multiple cloning site (MCS), where to insert the DNA fragment, flanked by sereval recognition sites for restriction endonucleases; and one or more genetic markers that allows the identification of cells that carry the cloning vector-insert DNA contruct[73].

In this study, the cloning vector used was pGEM-T[2] (Promega GmbH, Germany). This plasmid allows the direct cloning of PCR products with 5'-A residues at the extremities as result of the 5'→3' exonuclease activity of some DNA polymerases. With this enrichment of the 5' extremities, the DNA fragments amplified by PCR, are complementary to the 3'-T extremities of the cloning vector. This greatly improves ligation efficiency by the enzyme T4 DNA ligase. Another characteristic of this plasmid is that the MCS is located within the α-peptide coding region of the enzyme β-galactosidase (lacZ). When the DNA fragment is successfully ligated to the vector, this coding region is blocked and α-peptide is inactivated. As a result, the transformed cells that are grown in the presence of the *lac* operon, isopropyl-β-D-thiogalactopyranidose (IPTG) and the substrate 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal), are not able to synthesize the α-peptide, enabling them to hydrolyze the X-Gal in the growth media. Thus the colonies appear white[73].

The ligation reaction was carried out according to the manufacturer's instructions, with a fragment excess of 6:1, to increase ligation success. Ligation mix was carried out as described in table 3.4.

[2] *Vide* anex for plasmid vector map and reference points

Table 3. 4: Components (volumes in μl) of the reaction mix for the ligation of the Ngn3 fragments, amplified by PCR and purified, in the plasmid pGEM-T (Promega).

	Volume (μl)
2x Rapid Ligation Buffer	5.0
pGEM-T	1.0
Insert (Ngn3 fragment)	1.5
T4 DNA ligase	1.0
Water	1.5
Final volume	10.0

Reaction mix was incubated for 1 hour at room temperature and then used for the transformation of the competent cells.

3.2.2.4 Transformation of competent cells

The process of introducing DNA into a bacteria cell is named Transformation and the cells that are able of taking this new genetic material are said to be competent. Bacteria competence is due to altered cell walls that do not represent an efficient barrier against DNA uptake[73].

The bacterial cells used in this study were from the strain *E.coli* DH10B (Invitrogen, Germany), previously treated for competence by the calcium chloride ($CaCl_2$) method[3] and aliquoted by a member of the Group of Epigenetic Regulation of Transcription in Mouse Development and Disease.

For bacterial transformation, 5 μl of the ligation mix (section 3.2.2.3) were added to 100 μl of *E.coli* DH10B cells and incubated on ice for 10 min. Cells undertake a heat shock with a 1,5 min incubation at 42ºC and placed on ice. Then, 1 ml of Luria-Bertani (LB) broth were added to the cell suspension and incubated at room temperature for 1 hour. Bacteria were plated in LB agar plates with ampicilin (50μg/ml) supplemented with IPTG (1mM) and X-Gal (20μg/ml) and incubated at 37ºC overnight (O/N)

[3] *Vide* anex for the protocol on $CaCl_2$ competence method

Following incubation, plates were analyzed and the number of white colonies (with insert) and blue colonies (without insert) evaluated.

3.2.2.5 Miniprep

In order to extract the pDNA from the recombinant bacteria, the method of alkaline lysis was applied. In this method an alkaline solution (sodium hydroxide, NaOH) and a detergent (sodium dodecyl sulphate, SDS) are added to the cells to disrupt the cell walls and dissolve it, thus releasing the genetic material to the solution.

To purify the pDNA with the cloned Ngn3 fragment, 4 white colonies, from section 3.2.2.4, were isolated and transferred from the plate. Each colony was inoculated into 3 ml of LB broth, supplemented with ampicilin (50μg/ml) and incubated O/N at 37ºC, with agitation.

Cultures were transferred to 2 ml centrifuge tubes, centrifuged at room temperature for 5 min. The supernatant was discarded and the pellet resuspended in 100 μl of 50mM Tris-HCl, pH8/10mM EDTA. To the resuspended pellet, 100 μl of 0.2M NaOH/SDS 1% were added and the mix incubated at room temperature. To neutralize the solution, 150 μl of potassium acetate (KAc) 3M, pH 5,5 were added. The solution was mixed by inversion until flocculation. Tubes were centrifuged at maximum speed for 10 min, at room temperature. The supernatant was transferred to a fresh tube, 0.7V of isopropanol (Carl Roth GmbH, Germany) were added and mixed by inversion. Tubes were centrifuged at maximum speed for 25 min, at room temperature. The supernatant was then removed and discarded and the pellet washed. Pellet washing was carried out twice with the addition of ethanol (Carl Roth GmbH, Germany) 70% and centrifuged at maximum speed for 5 min, at room temperature. Supernatant was discarded, the pellet allowed to dry and resuspended in miliQ water.

Once resuspended the samples were quantified by spectrophotometry with Nanodrop and stored until further use.

The purified pDNA with the Ngn3 fragment inserted were sequenced at GATC Biotech (Germany) facilities, using SP6 primers to confirm their identity and orientation.

3.2.2.6 Midiprep

Once the identity of the cloned fragment was confirmed, a larger scale preparation of plasmid DNA was required. To achieve this goal, midiprep was performed. This protocol was used to prepare pDNA with the Ngn3 fragment and also for the other genes, mentioned on section 3.2.2.

To prepare the midipreps for the genes Muc4, Susd4 Igfbp5, Dusp1, Btg2, Ascl2, and Ngn3, the peqGOLD XChange Plasmid Midi kit (PeqLab Biotechnologie, Germany) was used according to the manufacturer's instructions. This protocol is based in the alkaline lysis method but it has the binding of the pDNA to an anion exchange resin under specific low-salt and pH conditions. Briefly, the transformed *E.coli* were inoculated in 100 ml of LB broth, incubated O/N at 37ºC with agitation (200 rpm). Cell suspension was transferred to 50 ml Falcon tubes and centrifuged for 10 min at 4300 rpm, at 4ºC. Supernatant was discarded and the pellet resuspended with 4 ml of Solution I/RNase A (provided by the kit). This was followed with the addition of 4 ml of solution II (from the kit) and all components were mixed by inversion. Mix was incubated for 5 min at room temperature and 4 ml of solution III (provided in the kit) precooled were added to the suspension the solution was mixed by inversion and centrifuged for 20 min at 4ºC at 4300 rpm. Resin columns were assembled and 2,5 ml of Buffer EQ (from the kit) were added to the columns in order to equilibrate them. The flow through was discarded and once the columns were equilibrated, the supernatant was transferred to the columns. The solution was allowed to pass through the resin by gravity and once all the liquid had pass, the column was washed with 10ml of DNA wash buffer and the flow through discarded. The columns were transferred to fresh tubes and 5 ml of Elution buffer (provided in the kit) were added to release pDNA from the resin. Columns were discarded and 3,5 ml of isopropanol were added to the flow through. Solution was mixed and centrifuged for 30 min at 4300 rpm at 4ºC. Supernatant was discarded, 10 ml of 70% ethanol were added to wash the pellet and the tubes centrifuged for 5 min at 4ºC

at 4300 rpm. Supernatant was discarded and pellet allowed to dry at room temperature until no ethanol drops were visible. Plasmid DNA was resuspended in miliQ water, quantified with the Nanodrop and stored at -20ºC until further use.

3.2.2.7 pDNA digestion

One of the steps prior to the synthesis of the *in-situ* probes is the linearization of the pDNA. Linearization will result in a site for the synthesis to end, as no stop codon is present in the pDNA. For this reason, enzymes were chosen with three criteria present: (1) cleave the construct only once, (2) cleave only in the MCS and (3) restrict the cleaving site to one of the insert flanks so that the cRNA would have the correct orientation (to pair with the mRNA expressed in the cells). Fragment orientation was determined by blasting the sequencing data against a data base (that confirmed that the gene Susd4 had not been cloned, and was not considered from this step forward).

Linearization of the vectors with fragments for Id2, Rik, Notch2, Zfp36, and Shh were previously prepared by a member of the laboratory.

Digestion reaction mix was as described in table 3.5.

Table 3. 5: Reaction mix for the restriction digestion of the pDNA for different genes. Digestion reactions were carried out with 10 mg of pDNA except for those marked with * (maximum amount of pDNA for the reaction volume of 100µl were the values presented; x corresponds to the volume of water needed to make 100µl of total volume; restriction enzymes, buffer and BSA from New England Biolabs).

	Muc4	Dusp1	Igfbp5	Btg2	Ascl2	Ngn3
Buffer 10X	10	10	10	10	10	10
BSA 100X	1	1	1	1	1	1
Enzyme	NdeI	NotI	NcoI-HF	NotI	NcoI-HF	NcoI-HF
Enzyme volume	4	4	4	4	4	4
DNA	10 µg	10 µg	10 µg	6 µg*	10 µg	4,9 µg*
Water	*x*	*x*	*x*	-	*x*	-

Tubes were spin down and incubated at 37ºC for 4 hours, after which a phenol:chloroform purification was performed. With this process restriction enzymes were removed from the solution retrieving the linearized pDNA with a high purity. Removal of the enzymes was possible due to the fact that phenol and chloroform denaturate proteins. Proteins then fall into the interphase between the organic and aqueous phases. By retrieving the aqueous phase, the nucleic acids are collected.

To carry out the phenol:chloroform purification, 200 µl of water were added, as well as 30 µl of NaCl 5 M and 300 µl of UltraPure Phenol:Chloroform:Isoamyl Alcohol (Invitrogen, Germany). Mix was gently vortexed and centrifuged for 2 min and room temperature at 13000 rpm. Upper phase was transferred to a fresh tube, 200 µl of chloroform (Carl Roth GmbH, Germany) were added and the mix was gently vortex and centrifuged for 2 min and room temperature at 13000 rpm. Upper phase was collected and 2.5V of 100% ethanol (Carl Roth GmbH, Germany) were added. Mix was vortexed and centrifuged for 25 min at 4ºC at 13000 rpm. Supernatant was removed and the pellets washed with ethanol 70%, centrifuged for 1 min at 13000 rpm at 4ºC. Supernatant was discarded and pellets allowed to dry at room temperature. When the ethanol was all evaporated, pellets were resuspended in 20 µl of miliQ water and quantified in Nanodrop.

Samples were stored at -20ºC until used for the synthesis of the RNA probes.

3.2.2.8 Reverse transcription and probe purification

Determination of the fragment orientation prior to linearization is important to make sure that the antisense strand of the insert will be synthesized. Otherwise the probe will not hybridize with the mRNA in the tissue and the *in situ* will not be successful.

The linearized plasmids and purified as described in the previous section were used to generate antisense DIG-labeled RNA-probes. They were generated with T7 RNA polymerase (Promega, Germany), if insert orientation was 3'→5' in the vector, or SP6 RNA polymerase (Promega, Germany), if the fragment presented the opposite orientation.

The synthesis of RNA labeled probes was as described in table 3.6.

Table 3. 6: Reaction mix for the synthesis of RNA DIG-labeled probes for ISH of slides. Synthesis of the RNA probes was carried out with T7 RNA polymerase or SP6 RNA polymerase, according to the fragment orientation. Probes were labeled with DIG using DIG labeling mix (Roche, Germany) and RNase Out (Invitrogen, Germany) to inhibit RNases activity.

Components	Volume (μl)
5X Transcription Buffer	4
DIG-labeling mix	2
Linearized pDNA	1
DTT (100 mM)	2
RNA Polymerase (T7/SP6)	2
RNase Out	0.5
Water	8.5

Mix was incubated for 2 hours at 37ºC. Once the synthesis was finished, 30 μl of miliQ water was added to each sample, and the probes purified using Illustra MicroSpin G-50 Columns (GE Healthcare, Germany), according to the manufacturer's instruction. Briefly, the columns were prepared by resuspending the resin with vortexing, the bottom closure was twisted off and the column was placed in the collection tube. Assembled tubes were centrifuged for 1 min at 2500 rpm and the flow through discarded. Columns were transferred to fresh RNase-free tubes, the mix with the newly synthesized probes were loaded into the columns and were centrifuged for 1 min at 2500 rpm. Flow through was recovered, samples quantified and stored at -80ºC until further use.

3.2.2.9 *In-situ* Hybridization of slides

Sections were rehydrated by an initial immersion for 5 min in Xylol (Carl Roth GmbH, Germany) twice, followed by a descending series of ethanol dilutions (twice in 100% for 5min, 70%, 50% and 30% for 2 min each) ending with a 5 min wash in PBS 1X.

Tissues were fixed with a 15 min immersion in 4% FA (Sigma-Aldrich, Germany) in PBS and washed in PBS 1X, twice for 5 min. Once the slides were washed, they were placed in a 6% H_2O_2 (Sigma-Aldrich, Germany) solution for 15 min, followed by 3 washes in PBS 1X for 2 min each. In order to increase the accessibility of the mRNA to the probes, tissues underwent a 10 min digestion with Proteinase K (PeqLab Biotechnologie, Germany) (10 μg/ml), a 15 min immersion in 4% FA in PBS for fixation and 2 washes in PBS 1X for 5 min each. For tissue acetylation, slides were immersed for 2 min in 100 mM Tris-HCl, pH 7,5[4], for 10 min in 0,25% Acetid Anhydride (Sigma-Aldrich, Germany) in 100 mM Tris-HCl, pH 7,5 and washed twice for 2 min in 2X SSC, pH 5. Tissues were dehydrated with immersion in a increasing series of ethanol dilutions (2 min in 30%, 50% and 70% each and twice for 5 min in 100%). Slides were allowed do dry for approximately 30 min prior to hybridization.

The hybridization buffer[5] was pre-warmed to 63ºC. RNA probe (in a final volume of 10 μl) was denaturated at 80ºC for 5-10 min on a ThermoCycler. The denaturated probe was added to the hybridization buffer (1μl of probe per ml of hybridization buffer). Once the hybridization solution was prepared, 100 μl were placed on each slide, with the borders previously defined using a Wax Pen. Slides with the hybridization solution were covered with a cover slip 24x60 mm #1,5 (Gerhard Menzel GmbH, Germany) and stored in a humid chamber in the dark at 63ºC overnight.

Following the incubation at 63ºC the slides were washed, at 60ºC with rocking, twice for 15 min in 2X SSC[6] in 50% Formamide (Sigma-Aldrich, Germany), twice for 15 min with 1X SSC and for 30 min with 0,2X SSC. Two final washes were performed with TBS at room temperature, 5 min each, followed by an incubation for 60 min with the Blocking Solution[7] in the humid chamber at room temperature. Antibody binding with Anti-DIG-AP, Fab fragments (Roche Diagnostics GmbH, Germany) (1:3000) occurred overnight at 4ºC in the humid chamber.

Following antibody binding, the slides underwent 6 washes of 30 min with TBSX[8] and 2 washes of 10 min with NTMT[9], all at room temperature with rocking. This was

[4] *Vide* Annex for composition of Tris-HCl
[5] *Vide* Annex for Hybridization Buffer composition
[6] *Vide* Annex for SSC composition
[7] *Vide* Annex for Blocking solution composition
[8] *Vide* Annex for TBSX composition
[9] *Vide* Annex for NTMT composition

followed by the staining step where 700 µl of staining solution[10] was added to each slide and incubated at room temperature in the dark in a humid chamber until color developed.

Staining reactions were terminated by immersing the slides in PBS 1X.

After staining, the tissues on the slides were dehydrated through an increasing series of ethanol (30%, 50%, 70% and 100% twice) 3 minutes each. Slides were transferred to Xylol (twice, 5 minutes each). Sections were mounted with Roti-Mount (Carl Roth GmbH, Germany) for permanent preparation and allowed to dry O/N under the fume cupboard. Slides were observed and photographed under the light microscope Leica DM2500.

3.3 Isolation of intestinal cells

This study focus on the changes in gene expression in the ISC and the possible mechanisms involved in the expression control. To be able to do so, it was required to isolate cell population of interest from all the different cells present in the tissue. This was performed by releasing cells from the tissue matrix into a cell suspension. Cells from the target population were isolated by FACS through specific markers used for cell determination.

The next sections will describe the methods through which samples were prepared for FACS.

3.3.1 Isolation of adult ISC

Adult stem cells were isolated from males of the Lgr5-EGFP-ires-creERT2 strain. This strain is a plus for this project as it takes advantage of the fact that adult ISC express the protein Lgr5. Through genetic manipulation, this strain expresses EGFP under the Lgr5

[10] *Vide* Annex for Staining solution composition

promoter, thus making it easier to isolate these cells from all the cells found in the small intestine.

After the small intestine was isolated, washed and sectioned as described in section 3.1, the fragments were transferred into a 50 ml Falcon. Tissue sections were washed with PBS 1X by inversion, until the solution was not cloudy. Solution was removed and fresh PBS 1X was added, supplemented with 5 mM EDTA. Tissue was incubated for 5 min at room temperature in a Table-shaker see-saw rocker SSL4 (Stuart Scientific, GB) at 40 rpm. After incubation the solution was mixed with gentle inversion. Solution was observed under the binocular for the presence of villus and/or crypts. Solutions rich in villus debries were discarded and washes were repeated with fresh PBS 1X/5 mM EDTA until the supernatant was poor in villus debris and begined to to be enriched in crypts. At this moment, aproximately 25 ml of PBS 1X were added to the tissue sections and the mix was shaked vigurously (≈ 10 harsh inversions). Supernatant was collected to a fresh tube and washes and supernatant recovery were repeated until no more crypts were present in the supernatant. Solutions contaning the crypts were centrifuged at 200g for 3 min and at 4ºC. Supernatant was discarded, pellet resuspended in 10 ml of PBS 1X and centrifuged again at 200g for 3 min and at 4ºC. Supernatant was removed, pellet resuspended in 10ml PBS 1X and transferred to a fresh tube. To dissosiate the cells making up the crypts, solution was digested with 2 μl dispase (100 mg/ml) (Sigma Aldrich, Germany) and 10 μl collagenase (50 mg/ml) at 37ºC for 30 min. During incubation, the suspension was pipeted to disolve any smeary clouds. Cells were observed for single events in a Cell counting device (Neubauer improved) (Carl Roth GmbH & Co., Germany). Cell suspension was centrifuged at 200g for 3 min. Supernatant was removed, pellet was resuspended in 1ml PBS 1X/2,5 mM EDTA and the suspension was rinse through a 30 μm Cell strainer cap into a 5 ml tube (BD Biosciences, Germany).

Cell suspensions were taken to the IMB cytometry facilities for cell sorting.

3.3.2 Isolation of embryonic ISC

After the embryos were isolated, as described in section 3.1, they were dissected under a Leica M80 to isolate the small intestine for further processing. The small intestines (pooled from several embryos due to the small size, thus low amount of cells) were placed in a tube with 500 µl of PBS 1X and digested with 5 µl of collagenase (50mg/ml) and incubated at 37ºC for approximately 15 min on a Thermomixer compact (Eppendorf AG, Germany), 800 rpm. During incubation the suspension was pipetted do dissolve smeary clouds. Cells were observed for single events in a Cell counting device and PBS 1X was added up to 1 ml. Cells were centrifuged at 2500 rpm, for 3 min at 4ºC. Supernatant was removed and pellet resuspended in 1 ml of PBS 1X and cells were centrifuged again at 2500 rpm, for 3 min at 4ºC. Supernatant was discarded and pellet resuspended in 100 µl of 2% Fetal Bovine Serum (FBS) (PAA Laboratories, Germany). To the cell suspension the following antibodies were added: 2,5 µl anti-CD31 (BD Biosciences, Germany) and 2,5 µl anti-EpCAM-PerCP-e710 (eBioscience, Germany) to ISC isolation of stage e12.5. For the isolation of ISC from stage e14.5, 2,5 µl anti-CD45-PE (BD Biosciences, Germany) were added together with EpCAM and CD31 as previously described. Cell suspension was resuspended with a pipette every 15 min and incubated in the dark for 30 min, at 4ºC. After incubation 900 µl of PBS 1X were added and centrifuged at 2500 rpm for 3 min at 4ºC. Washing was repetead and cells were resuspended in 500 µl PBS 1X/2.5mM EDTA. Cell suspension was rinsed through a 30 µm Cell strainer cap into a 5 ml tube prior to be taken to IMB cytometry facilities for cell sorting.

3.3.3 Isolation of adult enterocytes

Following small intestine isolation, washing and sectioning as described in section 3.1, sections were transferred into a 50 ml Falcon. Tissue was washed with PBS 1X by inversion, repeated until the solution was not cloudy. New washes were carried out with fresh PBS 1X supplemented with 5 mM EDTA. Tissue was incubated in for 5 min at

room temperature in a Table-shaker see-saw rocker SSL4 at 40 rpm. Solution was mixed by gentle inversion and observed under the binocular for the presence of villus and/or crypts. Solution rich in villus debries weas collected into a fresh 50 ml falcon tube and washes were repeated with fresh PBS 1X/5 mM EDTA until the supernatant was poor in villus debris. Villi were centrifuged for 3 min, 200g at room temperature. Supernatant was discarded and pellet resuspended in 20 ml of PBS 1X. Samples were centrifuged again for 3 min, 200g at room temperature. Supernantant was removed and pellet resuspended in 4,5 ml of PBS 1X. Villi suspention was distributed into fresh eppendorf tubes - 1 ml per tube and 0,5ml for negative control. To prepare cell suspention, samples and negative control were digested with 20 μg dispase and 50 μg collagenase. Digestion was incubated for 5 min at 37ºC. Cell suspension was centrifuged for 3 min at 2500 rpm, 4ºC. Supernatant was removed and pellet resuspended in 1 ml and 200 μl of PBS 1X, samples and negative control respectively and wash was repeated. Negative control was stored in ice until needed. Samples were resuspended in 200 μl of 2% FBS. The following antibodies were added: 2.5 μl anti-CD31, 2.5μl 45-PE and 2.5 μl anti-EpCAM-PerCP-e710. Cell suspension was resuspended with a pipette every 15 min and incubated in the dark for 30 min, at 4ºC. Following incubation, 900 μl of PBS 1X were added and centrifuged at 2500 rpm for 3 min at 4ºC. Washing was repetead and cells were resuspended in 500 μl PBS 1X/2.5mM EDTA. Cell suspension was rinsed through a 30 μm Cell strainer cap into a 5 ml tube prior to be taken to IMB cytometry facilities for cell sorting.

3.4 MBD-seq

MBD-seq is a widely technique to study methylation patterns of DNA. This is possible through the use of the proteins of the MBD family. As this proteins have the ability to bind specifically to the methylated cytosines in the DNA, they can isolate DNA fragments that are methylated from the remaining residue.

This experiment was carried out by a member of the Epigenetic Regulation of Transcription in Mouse Development and Disease group using the MethylMiner™ Methylated DNA enrichment kit (Invitrogen, Germany) according to the manufacturer's instructions. Samples used were ISC isolated by FACS as described in sections 3.3.1 and 3.3.2 for adult and embryo (e12.5) cells, respectively.

3.5 Bisulfite sequencing

One strategy to determine the methylation pattern of a given DNA sequence is the use of sodium bisulfate. DNA treatment with sodium bisulfate will deaminate unmethylated cytosine residues into Uracil (U), leaving the remaining C untouched. Thus, this treatment will give rise to different DNA sequences according to the methylation states of such sequences.

In order to validate the data from MBD-seq and have a more precise image of the methylation pattern, a few genes were chosen to undergo this treatment. The genes that were chosed were selected based on the MBD-seq data from adult and e12.5 ISC, in order for the genes to fall into the following groups: unmethylated in both stages, methylated in adult and unmethylated in embryo, methylated at e12.5 and unmethylated in adult, and methylated in both stages.

With the information about methylation in 2 distinct development stages already known, and to be confirmed by bisulfite sequencing, this was considered and opportunity to gather information about possible changes during development. Thus, samples used in this techique were ISC at stages e12.5, e14.5 and adult ISC and also adult enterocytes (AE), 3 replicates per stage.

The genes that were selected for this study were those described in table 3.7

Table 3. 7: Genes selected for methylation pattern analysis with sodium bisulfite. Genes were selected according to their methylation status in both staged analyzed, based on the data from the MBD-seq (section 3.4)

		Adult			
		Unmethylated		Methylated	
E12.5	Unmethylated	Ndufaf3		Nrp2	Meis 1
		Id2		Shh	Grb10
		Foxa1		Fzd2	Mdk
	Methylated	Olfm4	Vdr	Oct4	Efnb2
		Elf3	Ephb3	Lgr5	

Bisulfite sequencing included gDNA extraction and purification, sodium bisulfate treatment with the Epitect® Bisulfite kit (Qiagen, Germany), amplification of target fragments and purification.

3.5.1 Extraction and purification of gDNA

Genomic DNA was isolated from sorted ISC from stages e12.5, e14.5, adult and AE, as described in sections 3.3.1, 3.3.2 and 3.3.3, respectively.

Extraction of gDNA was carried out as follows. To each tube were added 5M NaCl, 0.5M EDTA, 0.5M Tris pH8 and SDS 10% to the final concentrations of 100mM, 10mM, 100mM and 0.1%, respectively. Samples were sonicated twice for 20 sec, with 90 sec pause, at high power and then treated with 1:1000 RNase A (Qiagen, Germany). For RNA digestion, samples were incubated for 30 min at 37ºC. This was followed by Proteinase K digestion for 5 hours at 55ºC. The DNA was now released from the cells and found in a mix of nucleic acids and cell debris. Purification of gDNA was carried out by phenol:chlorophorm precipitation. Briefly, 0.1V of 5M NaCl and 1V of UltraPure Phenol:Chloroform:Isoamyl Alcohol were added to the tubes. Mix was gently vortexed and centrifuged for 2 min and room temperature at 13000 rpm. Upper phase was collected to a fresh tube, 1V of chloroform was added and the mix was gently vortex and centrifuged for 2 min and room temperature at 13000 rpm. Upper phase was

collected and 2.5V of 100% ethanol were added. Mix was vortexed and centrifuged for 25 min at 4ºC at 13000 rpm. Supernatant was removed and the pellets washed with ethanol 70%, centrifuged for 1 min at 13000 rpm at 4ºC. Supernatant was discarded and pellets allowed to dry at room temperature. When the ethanol was all evaporated, pellets were resuspended in 25 μl of RNase-free water. Samples were quantified with Nanodrop and stored at -20ºC until needed.

3.5.2 Sodium Bisulfite treatment

Conversion of gDNA unmethylated C to U, was carried out with sodium bisulfite treatment using the Epitect® Bisulfite kit (Qiagen, Germany) according to the manufacturer's instructions.

To carry out this experiment, 1 μg of extracted and purified gDNA of each sample was used. DNA was thawed and the reaction mix was prepared as described in table 3.8.

Table 3. 8: Reaction mix for the sodium bisulfite treatment of gDNA from the Epitec® Bisulfite kit. Bisulfite Mix and DNA protect buffer were provided in the kit. * Combined volume of DNA and water should not be over 20 μl.

Component	Volume per reaction (μl)
DNA solution	1 μg (maximum 20 μl)
Water	Variable*
Bisulfite Mix	85
DNA protect Buffer	35
Total volume	140

Following the preparation of the reaction mixes, they were vortex thoroughly to ensure that the solutions were homogeneous. Bisulfite conversion was carried out in a thermal cycler and programed according to table 3.9.

Table 3. 9: Thermal cycler conditions, temperatures and step durations, for the bisulfite conversion using the Epitec® Bisulfite kit, for a final reaction volume of 140 µl.

Step	Duration (min)	Temperature (ºC)
Denaturation	5	95
Incubation	25	60
Denaturation	5	95
Incubation	85	60
Denaturation	5	95
Incubation	175	60
Hold	Indefinite	20

Once the bisulfide conversion was completed, products underwent a series of cleanup steps. Following the reaction in the thermal cycler, samples were centrifuged and transferred to clean 1.5 ml microcentrifuge tubes. To each reaction mixture 560 µl of freshly prepared Buffer BL (provided in the kit) were added. Solution was vortex and spined down. The mixture was tranferred into the previously prepared columns. Columns and collection tubes were centrifuged for 1 min, maximum speed at room temperature. The supernatant was discarded and the columns placed back in the collection tudes. To each column 500 µl of Buffer BW were added. Columns were centrifuged for 1 min, maximum speed at room temperature, flow-through discarded and 500 µl of Buffer BD were added. Columns were incubated with Buffer BD for 15 min at room temperature and centrifuged for 1 min, maximum speed at room temperature. Flow-through was discaded and DNA bound to the column was washed twice by adding 500 µl of Buffer BW, centrifuging (1 min, maximum speed) and discarding the flow-through. Columns were transferred to freshcollection tubes and spined down to remove residual liquid. The spin columns were placed in fresh 1.5 ml microcentrifuge tubes and 20 µl of Buffer EB were added onto the center of the

membrane. Purified DNA was eluted with a 1 min centrifugation at maximum speed and room temperature.

Samples were stored at -20ºC until needed.

3.5.3 Target amplification

With the approach taken for the analisys of the methylation pattern, through bisulfite sequencing, it was necessary to focus in a few regions of the gDNA. In order to to so, specific primers were designed to amplify fragments of the genes selected for analysis (table 3.7).

3.5.3.1 Primer design

Primers were designed using MethPrimer online software (www.urogene.org/cgi-bin/methprimer/methprimer.cgi) that predicts the sequence after citosine conversion by sodium bisulfite treatment and present several options of specific primers that can be used for PCR.

General parameters for primer design used were as default with the exception of the fragment size that ranged from 300 to 500 bp, with 400 bp being set as optimal fragment size. These sizes were defined according to the limitations of sequencing analysis.

Primers designed were those presented on table 3.10.

Table 3. 10: Primer sequences of the specific primers for amplification of gene fragments of the selected genes for analysis of methylation pattern and expected size of amplicon (bp). Primers were designed with MethPrimer for amplicon size 400 ± 100 bp.

Gene		Primer sequences (5'→3')	Amplicon size (bp)
Ndufaf3	Fw	GGTTGGATAGGGGATGTTTATTATAG	468
	Rv	CACACCCTAACCAAATATTCATAAC	
Id2	Fw	TAAATGTTTGTAAAAGATTGGGAGAG	329
	Rv	CCCCTATACTCTAAAATCCAAATTTC	
Foxa1	Fw	GTTTAGAGTTTGGGGGTTGG	401
	Rv	AAATCATAAACCTCTTCCCCTATTAC	
Nrp2	Fw	GGGTTGAAGTTGGTGAAGTAAGTTA	379
	Rv	AAAACAATCCCAAAATAAAAAAACC	
Shh	Fw	GGAAGGGATGTAGTAATAATAGGTT	481
	Rv	ACTCTAAAACACCAAATACCAATTC	
Fzd2	Fw	TTAAAAGTTGTTTTGTGTTTTGGTT	409
	Rv	ATAAAAAATAAAAACCATTCCTCCC	
Meis1	Fw	ATTTTTGATGTTAATTGGTTGTTTG	371
	Rv	CTTAATAAAAATAACCCCTCACTCC	
Grb10	Fw	TTTTGGTTTTTAGATTTGTGAAATG	314
	Rv	CTCCCAAAAATAAATCTTACCTACC	
Mdk	Fw	TTTTAGGGGTAAAATTTGGGTTAG	327
	Rv	TTCTAAAAACAAAAAAAACCTTAAC	
Olfm4	Fw	GTGTGTATGAGAAGAGGTTTTTGAAT	306
	Rv	AACCAACTAACCTTAAACACACAAA	
Elf3	Fw	TTGAGTTTATTAGAAGAGTTGGAGGTT	429
	Rv	ACAAAATCCCAAAAAAACAAAACTA	
Ephb3	Fw	TGGTTTTTTAGGATTAAGGTGTTTG	445
	Rv	TCTCTCCCATATCCCTCTTTATAAAA	
Vdr	Fw	GTTTGTTTGTTTGTTTGTTTGTTTG	445
	Rv	CTTACCCTAATTACCAAAACCCTCT	
Oct4	Fw	GTTGTGAGTTATTTTGTGGTTGTTG	423
	Rv	AATACCCCAACAAATTCCTATCTACT	
Lgr5	Fw	GGGGAGGGAAGATAGATAGATAGATAG	401
	Rv	CCCCAAAACAAAAATTCTAAAATTAC	
Efbn2	Fw	TGTTTGTAGTTTTTAGTGGTATGATAA	413
	Rv	CTCCTTTACCCTATCTATCAAATTC	

Primers were resuspended according to the manufacturer's (Sigma-Aldrich, Germany) intructions for a final concentration of 100 μM using miliQ water.

3.5.3.2 PCR conditions

Conditions for the PCR reactions were optimized in order to set the best temperature of annealing (Ta) and the concentration of magnesium cloride ($MgCl_2$) for each pair of primers.

The choice of Ta is largely determined by the melting temperature of the pair of primers, estimated by the software. As the behaviour of new primers during a PCR reaction is not well known, they were tested in a gradient of temperatures, starting at a lower temperature – less specific – up to a higher temperature – more specific, thus reducing the probability of secondary products. In respect to the concentration of $MgCl_2$, it was tested simultaneously with the temperature gradiente using also highly stringent conditions, with 1.5 mM $MgCl_2$ as well as a more permissive concentration – 2 mM.

This optimization was also an opportunity to assess the primers that were design and establish those that were able to produce an amplicon.

In order to reduce the risk of misreading of the polymerase enzyme during the PCR reaction, thus leading to incorrect data, a high fidelity system, AccuPrime™ *Taq* DNA Polymerase System (Invitrogen, Germany), was used.

After optimization, the best conditions for amplifing the different gene fragments, in respect to $MgCl_2$ concentrations were as presented on table 3.11.

Table 3. 11: Conditions, in respect to MgCl2 concentrations, to be used for amplification of the different genes.

	Low $MgCl_2$ (1.5 mM)	High $MgCl_2$ (2mM)
Genes	Ephb3	Elf3
	Ndufaf3	Grb10
	Fzd2	Id2
	Lgr5	Vdr
	Oct4	
	Olfm4	

Thus, the final conditions for PCR mix with the different $MgCl_2$ concentrations are presented on table 3.12.

Table 3. 12: PCR reaction mix for amplification of gene fragments in $MgCl_2$ stringent and less stringent conditions (1.5 mM and 2 mM of $MgCl_2$, respectively) for a final volume of 50 µl.

Component	Low $MgCl_2$ (1.5 mM) Volume (µl)	High $MgCl_2$ (2mM) Volume (µl)
10X AccuPrime PCR Buffer II	5	5
Primer Mix (1µM)	10	10
AccuPrime™ *Taq* DNA Polymerase	0.85	0.85
DNA converted (1:5)	2	2
$MgCl_2$ (50 mM)	---	1.25
Water RNase-free	32.15	30.90
Total volume	50	50

During optimization, the Ta for each pair of primers were also determined. They were as presented in table 3.13.

Table 3. 13: Temperatures of annealing (Ta) for the amplification of the fragments of the different genes by PCR reactions.

Temperatures	52ºC	55ºC	56ºC	60ºC
Genes	Ephb3	Fzd2	Lgr5	Vdr
	Ndufaf3	Elf3	Oct4	
		Grb10	Olfm4	
			Id2	

Amplification reactions by PCR were carried out in a thermo cycler as follows (table 3.14).

Table 3. 14: Cycling conditions for the PCR reactions for amplification of fragments of the different target genes after sodium bisulfate treatment. Ta specific for each pair of primers are presented on table 3.13.

Temperature (ºC)	Duration	Number of cycles
95	5 min	
95	35 sec	4 cycles (with Ta decreasing 1ºC per cycle)
Ta+4	30 sec	
68	50 sec	
95	35 sec	37 cycles
Ta	30 sec	
68	50 sec	
68	7 min	
4	4 min	

Reaction products were analysed by electrophoresis in a 1.5% agarose gel stained with ethidium bromide. To 10 µl of PCR products, 2 µl of 6X Orange DNA Loading Dye (Fermentas) were added, and the final volume was loadded into the gel for amplicon visualization. This was carried out in a Chemidoc XRS system (Biorad, Germany).

When required, reamplification was performed under the same conditions as described in tables 3.12 and 3.14, using as template the PCR products from the first reaction.

4. Results

Throughout this study two main approaches were taken: morphological and molecular. The morphological techniques were used to identify the location where a given gene was expressed in the tissue of interest – small intestine in different developmental stages. With the molecular approach, it was aimed to validate and acquire a more detailed perspective of the changes in the methylation patterns identified by MBD-seq.

4.1 Sampling

The study of gene expression throughout the development of the small intestine in mice required the collection of adult small intestine and mouse embryos at different developmental stages. In Fig 4.1 mouse embryos at stages e12.5 and e14.5 (images A and B, respectively) as well as a small portion of adult small intestine (image C) rich in villus, are presented.

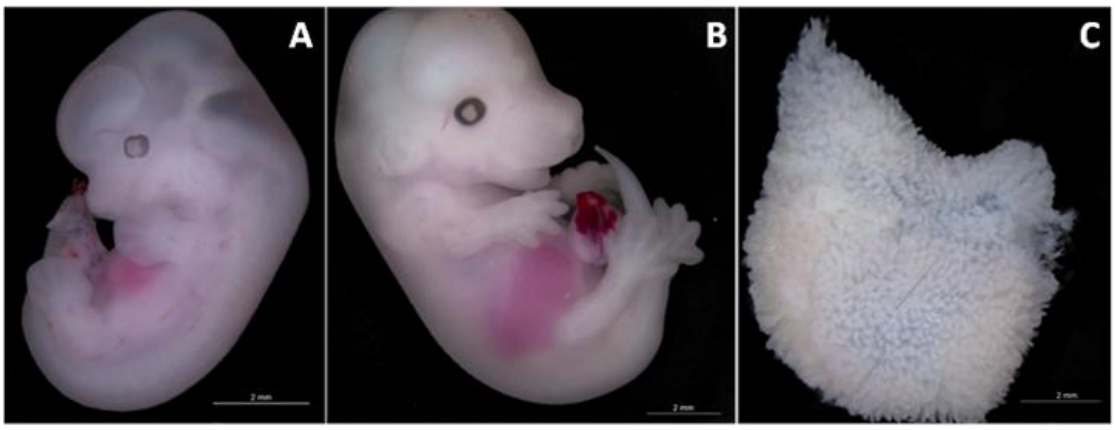


Fig 4. 1: Mouse embryos at developmental stages e12.5 (A) and e14.5 (B) and small section of adult small intestine (C).

4.2 *In-situ* hybridization

To gaininformation about the specific location of a given set of genes, the ISH technique was used. By targeting a gene's mRNA with complementary probes it was possible to locate where the target gene was expressed. Target genes were selected based in data from RNA-seq previously prepared of ISC from e12.5 and adult. Through this process an accurate estimation of transcript abundance was mapped[75] producing genome-wide transcriptional maps, in this case, of the ISCs at e12.5 and adult stages. Expression levels were determined by initially converting the isolated RNA into a cDNA library. Then, the reads were mapped against the mouse genome as reference, followed by counting the number of reads that aligned to the DNA sequence. Comparing which genes were expressed at the ISC in stage e12.5 and adult, was possible to observe that the genes expressed in mouse ISCs are different between the two developmental stages

In order to carry out the ISH on slides, several intermediate steps were carried out in order to prepare the probes and to ensure they were in the best conditions.

4.2.1 Fragment amplification for RNA probe preparation

To assess the success of the amplification reaction for the Ngn3 gene, 10 μl of loading buffer were added to the reaction volume and the total volume was loaded into a 1.5% agarose gel in 1X TAE, stained with EtBr (Fig 4.2).

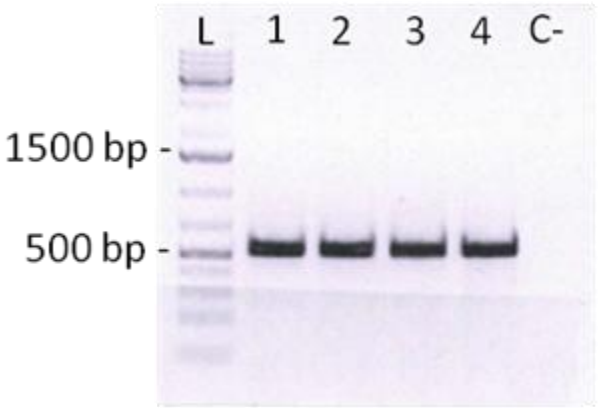


Fig 4. 2: Products of PCR reaction for the amplification of a fragment of Ngn3 on a 1.5% agarose gel in 1X TAE, stained with EtBr. Lanes 1 to 4 are different PCR products, specific for Ngn3, lane L is the DNA ladder Gene Ruler 1Kb Plus DNA and lane C- is the negative control for the PCR reaction. (Samples run from the highest to the lowest molecular weight).

The PCR reaction performed for the amplification of a Ngn3 fragment appeared to be specific. The PCR reactions produced one single band with the a size higher than 500 bp (expected sized of the fragment was 608 bp).

Contaminations were discarded as the negative control was clear.

PCR products were then processed (purified and cloned) to determine fragment's identity. This was achieved by sequencing at the GATC Biotech facilities, which confirmed that the cloned fragment was Ngn3. Following identity confirmation, a larger amount of pDNA was required and prepared. The pDNA was then linearized for further use in reverse transcription of the cRNA probes.

4.2.2 Reverse transcription

Once the reverse transcription was finished and the probes purified, they were loaded onto an agarose gel to assess the success of the reaction. Probes that had been previously prepared and were stored were also loaded into an agarose gel. This allowed to check for any signs of probe deterioration.

Probes were loaded into different gels. In the first gel cRNA probes for Ascl2, Muc4, Igfbp5, Btg2 and Dusp1 were loaded (Fig 4.3).

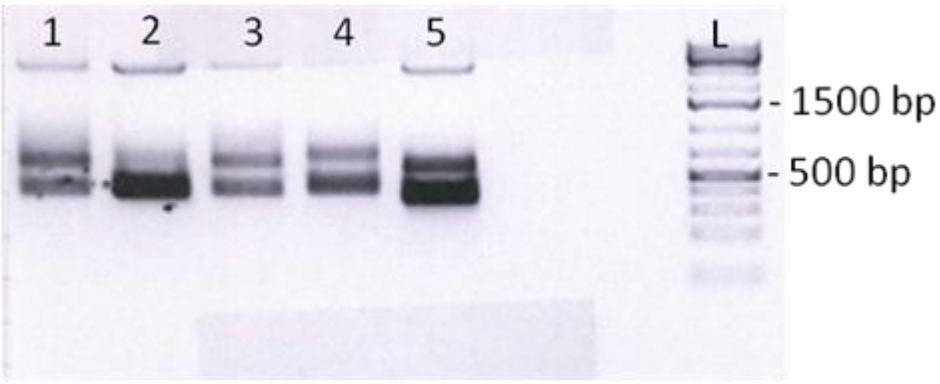


Fig 4. 3: Probes for the ISH method (5 µl of each probe) in a 1.5% agarose gel in 1X TAE, stained with EtBr. cRNA probes DIG-labeled were Ascl2 (lane 1), Muc4 (lane 2), Igfbp5 (lane 3), Btg2 (lane 4) and Dusp1 (lane 5). Lane L is the ladder Gene Ruler 1Kb Plus DNA. (Samples run from the highest to the lowest molecular weight).

From the observation of the gel, it was confirmed that the probes for Ascl2, Muc4, Igfbp5, Btg2 and Dusp1 were in good conditions and were used downstream in the ISH protocol.

Following the reverse transcription to prepare the probe for Nocht2, 5 µl of probe, DIG-labeled were loaded onto a 1% agarose gel in 1X TAE stained with EtBr (Fig 4.4) for quality analysis.

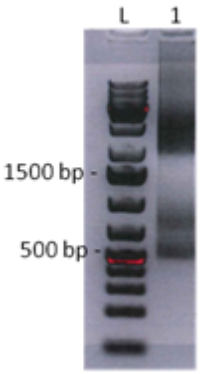


Fig 4. 4: Notch2 probe, DIG-labeled, with 5 µl of Orange DNA Loading dye in a 1% agarose gel stained with EtBr. Lane L is the ladder Gene Ruler 1Kb Plus DNA and lane 1 is the probe for Notch2. (Samples run from the highest to the lowest molecular weight).

Analyzing the gel from the electrophoresis on the cRNA probes for Notch2 (Fig 4.4) it was confirmed also that the probe was in good conditions and could be used for ISH on slides. The low sharpness of the bands was due to the low percentage of agarose.

The probe for Ngn3 was analyzed by electrophoresis by loading 5 μl of probe, DIG-labeled, onto a 1% agarose gel in 1X TAE stained with EtBr (Fig 4.5) for quality analysis.

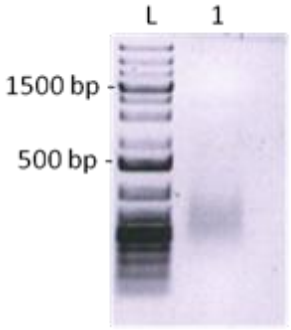


Fig 4. 5: Ngn3 probe, DIG-labeled, with 5 μl of Orange DNA Loading dye in a 1.5% agarose gel stained with EtBr. Lane L is the ladder Gene Ruler 1Kb Plus DNA and lane 1 is the probe for Notch2. (Samples run from the highest to the lowest molecular weight).

The Neurogenin3 (Ngn3) probe was also in proper conditions to be used in the ISH of slides. The faint smear on the gel (Fig 4.5) was due to a low concentration of probe and also to the low percentage of agarose of the gel.

4.2.3 ISH on slides

This task of the current study was performed in order to validate the RNA-seq data. It was also an opportunity to gather information about the changes in the gene's expression during the embryonic development. Thus gene expression was also analyzed in stages e13.5, e14.5 and e15.5.

Genes analysed include Ascl2 (Fig 4.6), Btg2 (Fig 4.7), Cps1 (Fig 4.8), Dusp1 (Fig 4.9), Hes1 (Fig 4.10), Igfbp5 (Fig 4.11), Kcne3 (Fig 4.12), Muc4 (Fig 4.13), Notch2 (Fig 4.14), Shh (Fig 4.15) and Ngn3 (Fig 4.16).

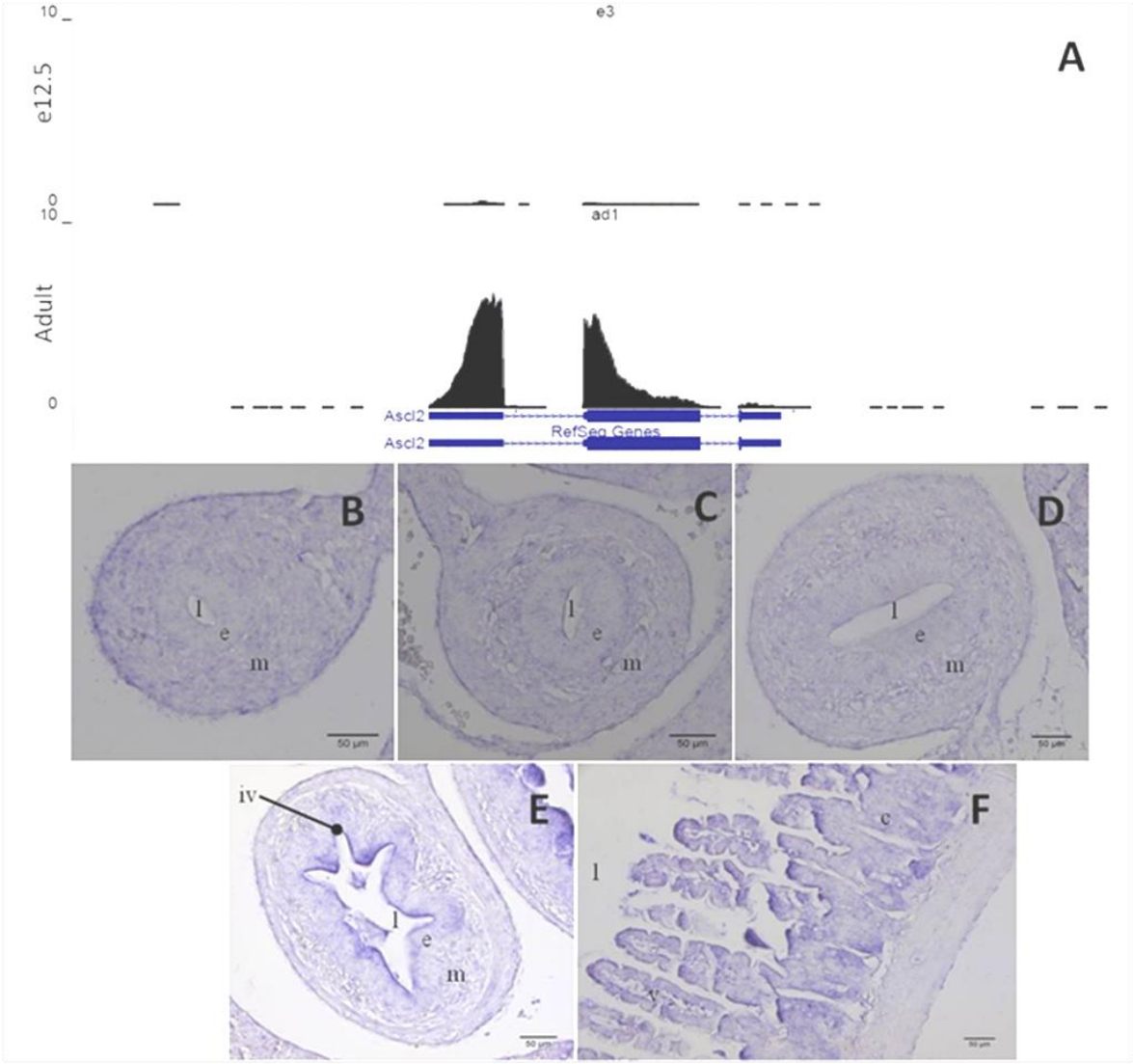


Fig 4. 6: Expression pattern of the gene Ascl2 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. ISH specific for the gene Ascl2 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus region, v – villi, c – crypt. Scale bar corresponds to 50μm.

In respect to the results on the expression of Ascl2 (Fig 4.6), it was noticed a higher expression on the adult ISC (Fig 4.6 A). Regarding the ISH targeting Ascl2 on sections of small intestine (Fig 4.6 B to F), no expression was detected in the sections from

embryos (Fig 4.6 B to E), since there are no regions with higher intensity of staining. In the sections of the adult small intestine (Fig 4.6 F) it appears that some regions of the crypts present a slight more intense staining.

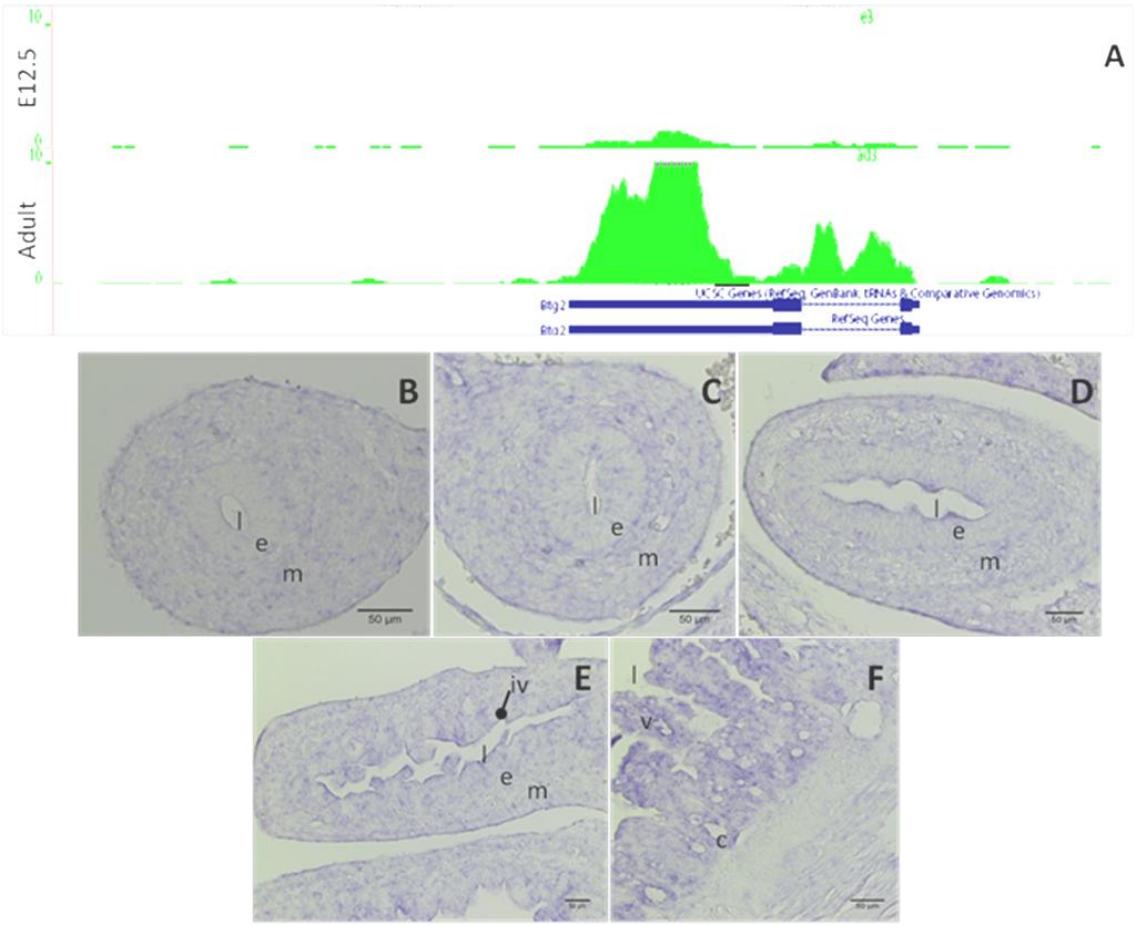


Fig 4. 7: Expression pattern of the gene Btg2 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. ISH specific for the gene Btg2 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus region, v – villi, c – crypt. Scale bar corresponds to 50µm.

From the RNA-seq data for the gene Btg2 (Fig 4.7 A) the expression of this gene is stronger in the adult ISCs when compared with the expression levels detected for the embryonic ISCs that were isolated. In the ISH study of the gene Btg2 (Fig 4.7 B to E) no increase in staining is detected. The absence of staining intensity in the sections from e12.5 (Fig 4.7 B) are in accordance with the data from the RNA-seq data. Nevertheless,

the staining in the sections from the adult tissue (Fig 4.7 F) show a stronger staining in the epidermal layer when compared with the underlying tissues of the small intestine.

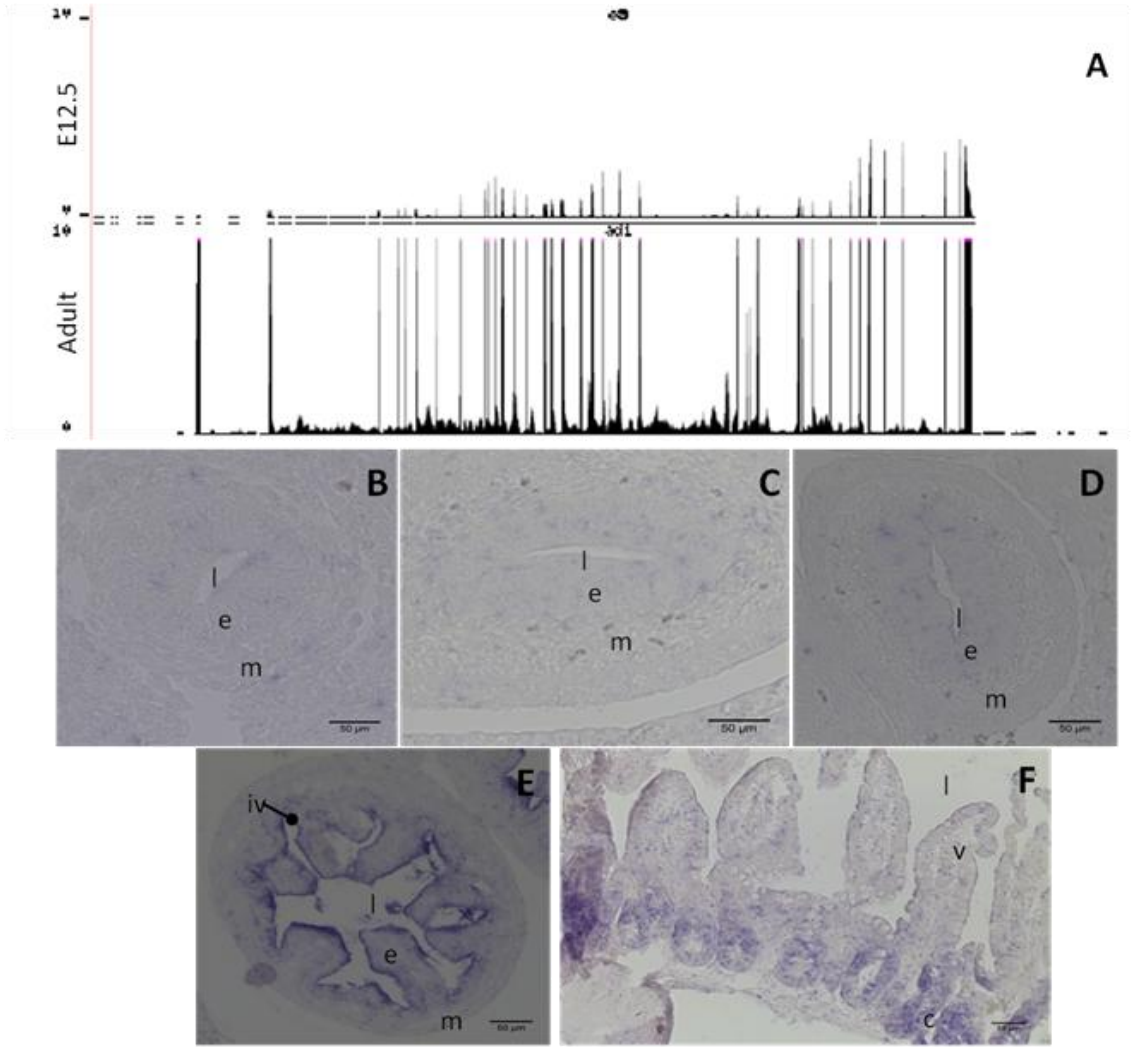


Fig 4. 8: Expression pattern of the gene Cps1 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. ISH specific for the gene Cps1 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50μm.

Observing the expression patterns of the gene Cps1 (Fig 4.8), it was perceived, from the RNA-seq data (Fig 4.8 A), that this gene had a higher expression in the adult ISC when compared with the ISC at stage e12.5. Data from RNA-seq are further supported with the ISH results (Fig 4.8 B to F). The results from the ISH method show, on adult small intestine (Fig 4.8F) a higher intensity of staining in the crypt. Comparing to the other

stages (Fig 4.8 B to E), no increasing on staining intensity is noted. Nevertheless, it is possible to note that are some spots in the intestinal endothelium in the stages e12.5 to e14.5 (Fig 4.8 B to D). In the interface between the lumen and the intestinal endothelium in the stage e15.5 (Fig 4.8 E) there is a strong unspecific staining not related to specific hybridization, but rather caused by the accumulation of pigment.

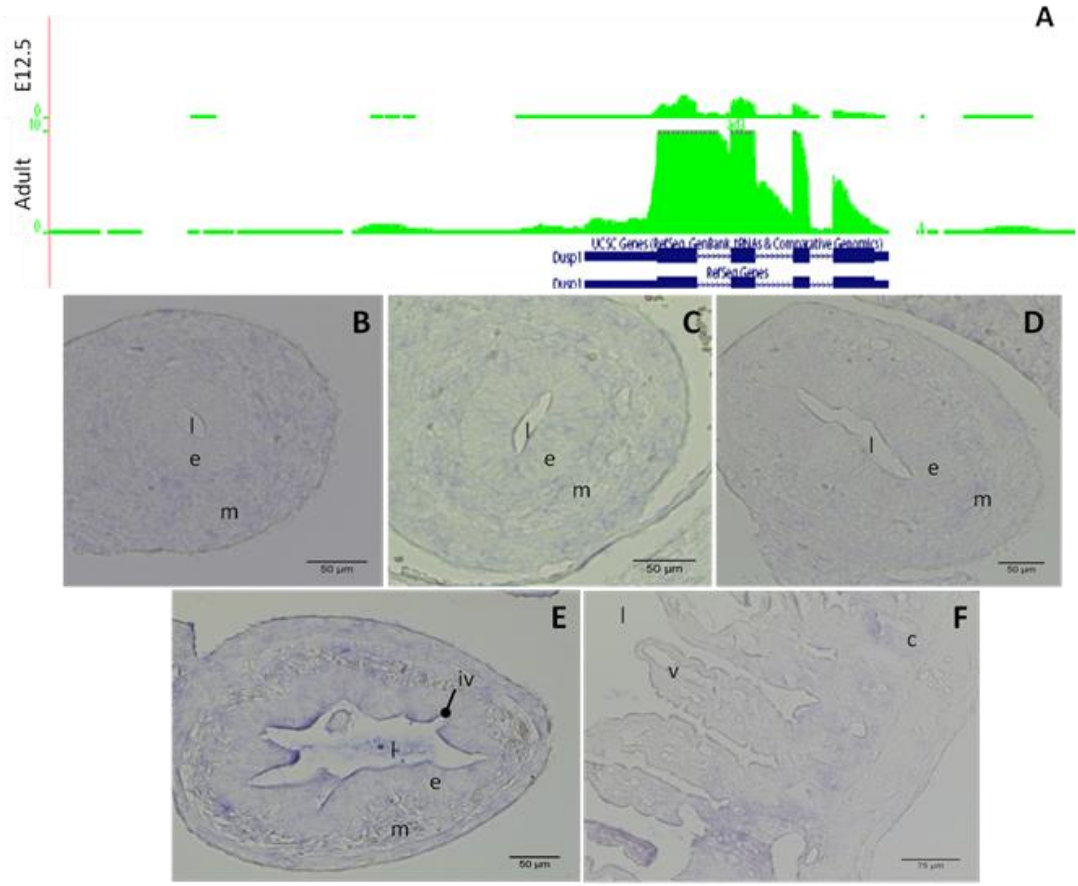


Fig 4. 9: Expression pattern of the gene Dusp1 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. ISH specific for the gene Dusp1 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50µm to all sections except F, where it corresponds to 75µm.

Observing the results from the RNA-seq and ISH on the gene Dusp1 (Fig 4.9), some differences can be noticed. In respect to the RNA-seq data (Fig 4.9 A), a strong expression of the gene was detected in the population of adult ISCs. Nevertheless, a very low level of expression of Dusp1 is detected in the embryonic samples from e12.5. The ISH (Fig 4.9 B to F), on the other hand, presented with no changes in the

expression of the gene (which would be seen as an increase in staining intensity if the amount of transcripts were higher) in the stage e12.5 (Fig 4.9 B) although an increase of staining intensity could be noticed in the sections of adult tissue (Fig 4.9 F), at the bottom of the crypts. Also no changes in the staining were detected in the sections from e13.5, e14.5 and e15.5 (Fig 4.9 C, D and E respectively) indicating that this gene is not transcribed during embryonic development.

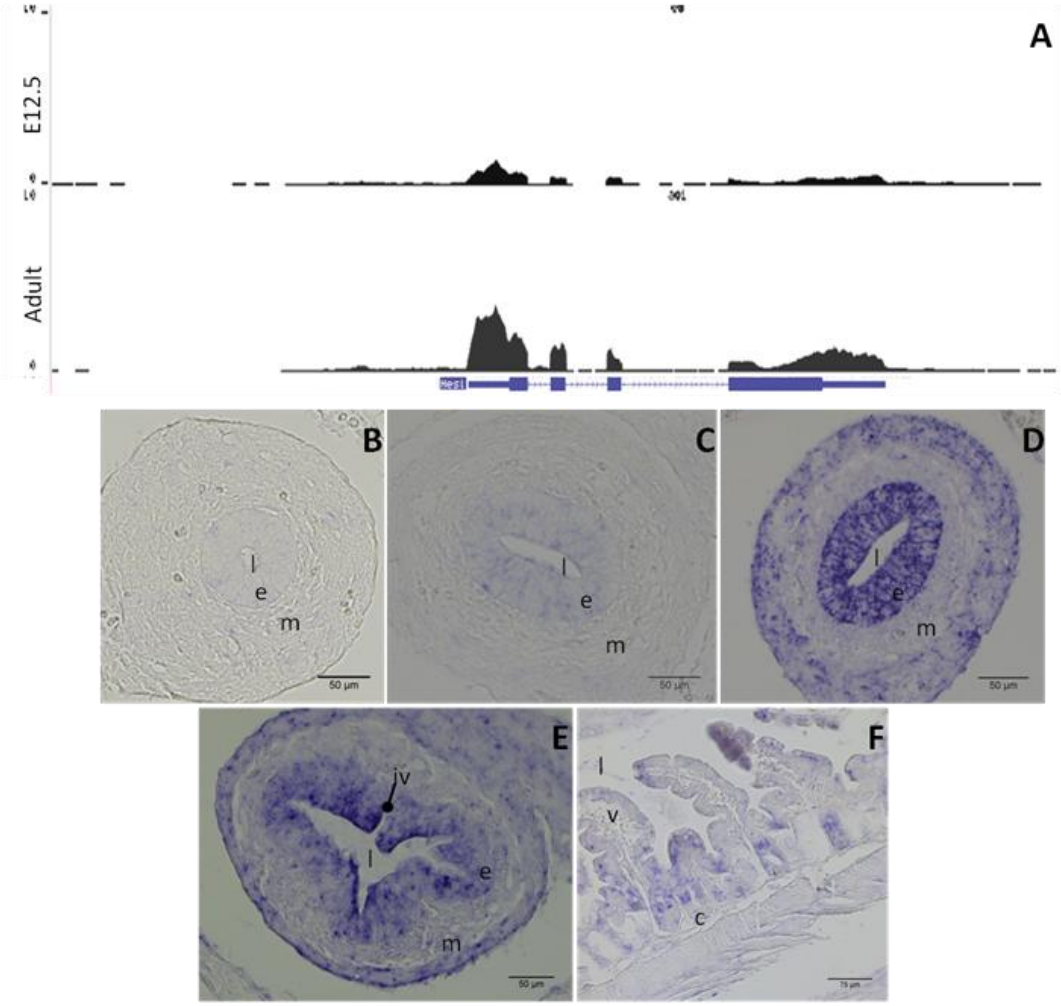


Fig 4. 10: Expression pattern of the gene Hes1 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. ISH specific for the gene Hes1 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50µm to all sections except F, where it corresponds to 75µm.

The expression analysis of Hes1 in the RNA-seq data (Fig 4.10 A) shows that this gene has a higher expression in the cells from the adult small intestine. The ISH sections

(Fig 4.10 B to F) confirms that expression pattern. In the section from the developmental stage e12.5 (Fig 4.10 B) no staining is detected relative to the target gene. As development continues (e13.5, e14.5 and e15.5) an increase in the staining intensity can be noticed (Fig 4.10 C to E). The expression of Hes1 continues in the adult small intestine (Fig 4.10 F), specifically in the intestinal crypts. This validates the RNA-seq data and also shows that expression of the gene Hes1 starts at early embryonic development and not only in the adult small intestine.

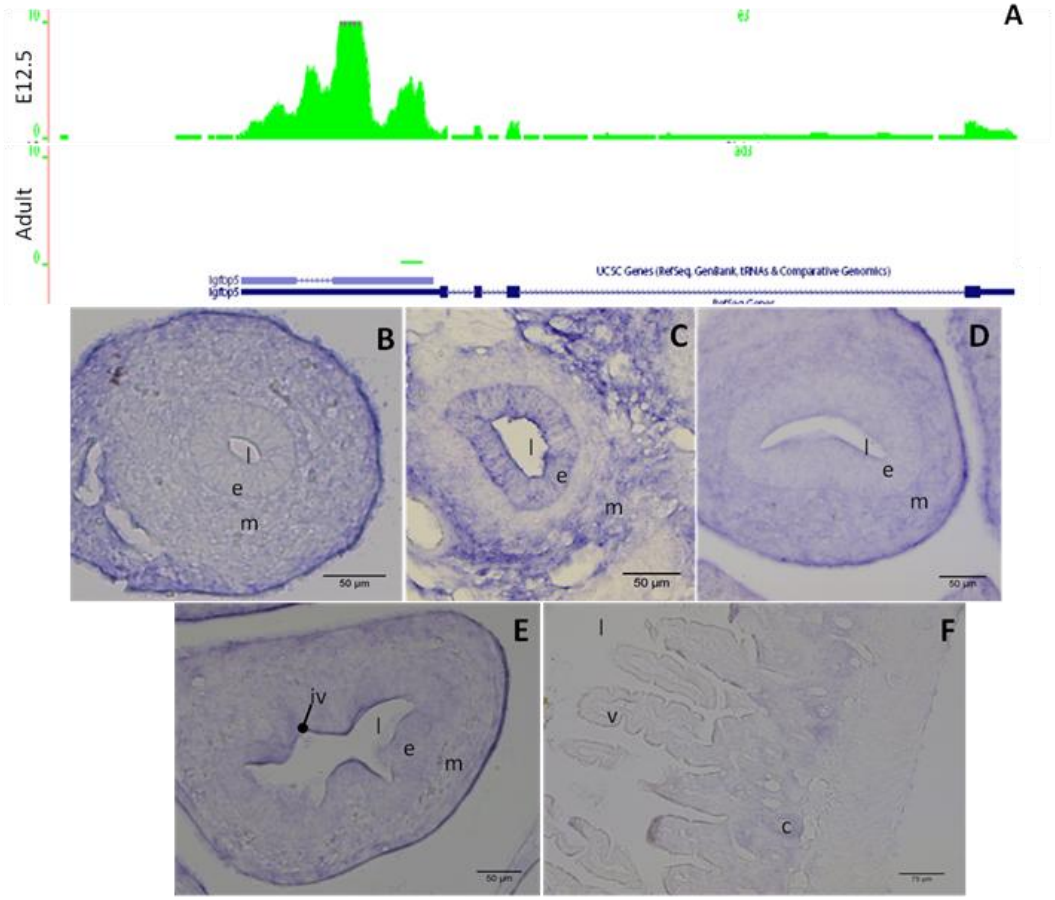


Fig 4. 11: Expression pattern of the gene Igfbp5 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. Note the exon represented with the solid box at the bottom. ISH specific for the gene Igfbp5 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50µm to all sections except F, where it corresponds to 75µm.

Analysis of the expression of the gene Igfbp5 was performed based on RNA-seq and ISH (Fig 4.11). From the RNA-seq (Fig 4.11 A), it can be observed that this gene is

expressed only at stage e12.5, as it was in these samples that a higher amount of transcripts was detected. The results from the ISH (Fig 4.11 B to F) show no differences in the intensity of the staining in the stages e12.5 and the adult ISC (Fig 4.11 B and F, respectively). Section from e13.5 show a signal in the endothelium (Fig 4.11 C) which may indicate that these gene is not expressed only at stage e12.5 (form RNA-seq data).

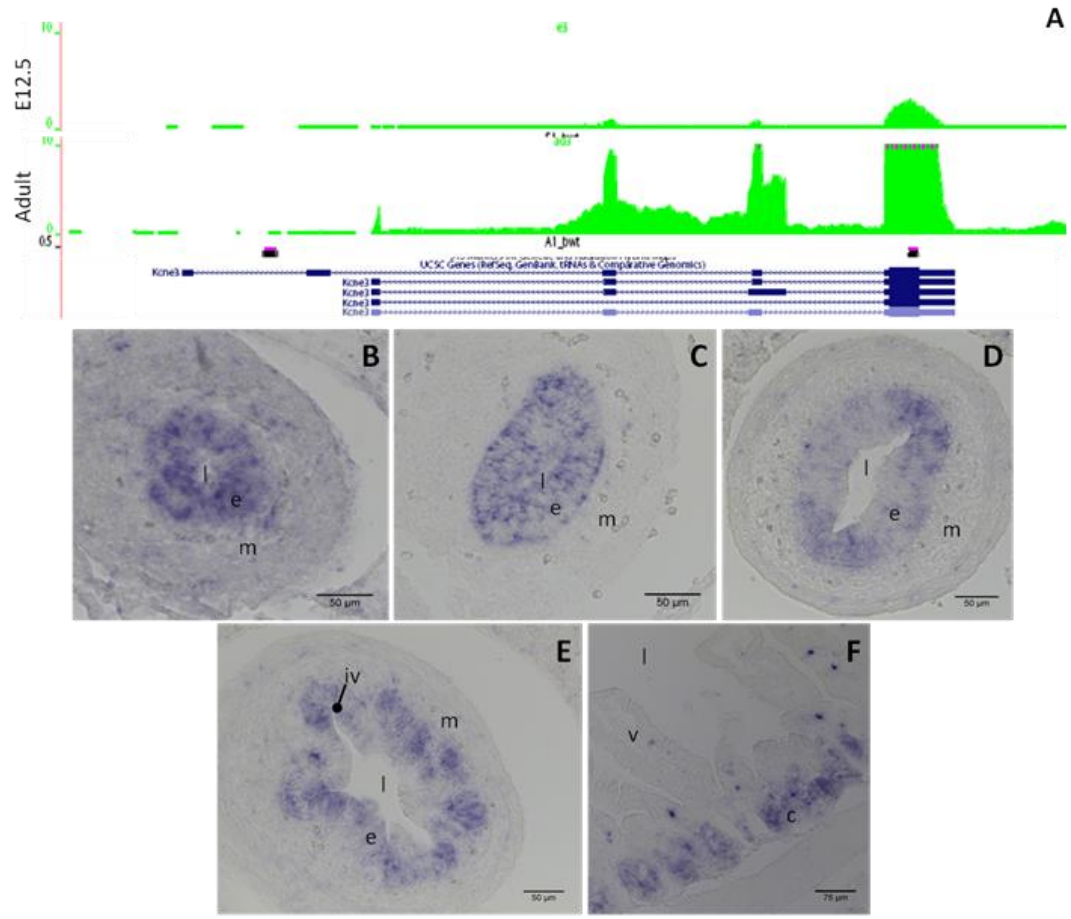


Fig 4. 12: Expression pattern of the gene Kcne3 resulting from RNA-seq (A) and ISH (B to F). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. Note the exon represented with the solid box at the bottom. ISH specific for the gene Kcne3 in fixed sections of mouse embryos e12.5 (B), e13.5 (C), e14.5 (D) and e15.5 (E) and adult mouse small intestine (F). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50µm to all sections except F, where it corresponds to 75µm.

The expression analysis of Kcne3 by RNA-seq (Fig 4.12 A) show that this gene is highly expressed in the population of adult ISCs isolated. This data also indicate that there is some expression in the embryonic endothelial cells (e12.5), although not as

strong as the detected in the adults. Staining for the gene Kcne3 shows its expression at e12.5 (Fig 4.12 B) spread throughout all endothelium, while in the adult sections (Fig 4.12 F), gene expression is detected at the bottom of the crypts. Regarding the intermediate embryonic stages e13.4, e14.5 and e15.5 (Fig 4.12 C, D and E, respectively) gene transcripts were also detected indicating that the gene is present in all developmental stages. Note that at stage e15.5 (Fig 4.12 E), the expression is detected in more localized region – the intervillus regions that start to appear at this stage.

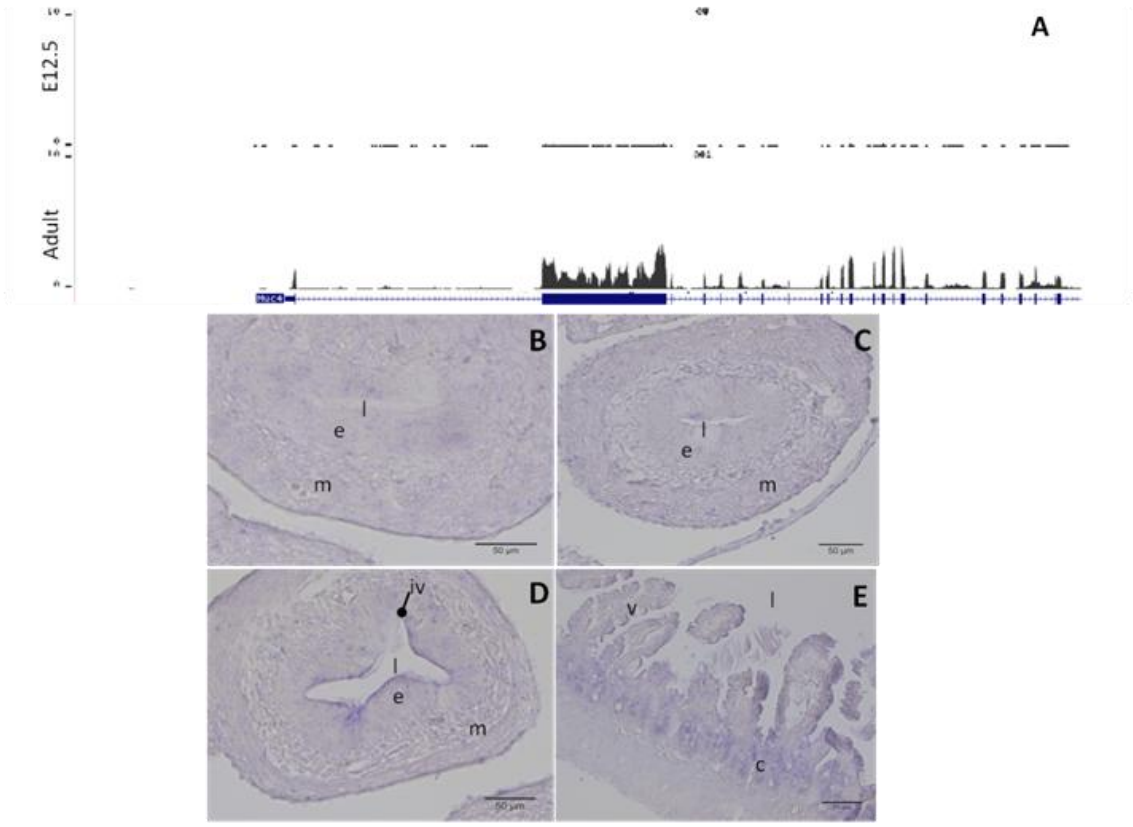


Fig 4. 13: Expression pattern of the gene Muc4 resulting from RNA-seq (A) and ISH (B to E). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. Note the exon represented with the solid box at the bottom. ISH specific for the gene Muc4 in fixed sections of mouse embryos e12.5 (B), e14.5 (C), e15.5 (D) and adult mouse small intestine (E). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50μm to all sections except E, where it corresponds to 75μm.

The data from the RNA-seq on the gene Muc4 (Fig 4.13 A) gives the indication that this gene has a higher expression in the adult ISC in comparison with e12.5. The complementary analysis on this gene by ISH (Fig 4.13 B to E) is in accordance with this

observation. Focusing on the section from the adult small intestine (Fig 4.13 E), a slight increase in the intensity of the staining can be noticed in the cryps. Changes in the intensity of the staining on the intestinal endothelium, when compared to the surrounding tissues, in the embryonic stages (Fig 4.13 B to D) present no changes. This could be indicative that the gene is not expressed at such early stages.

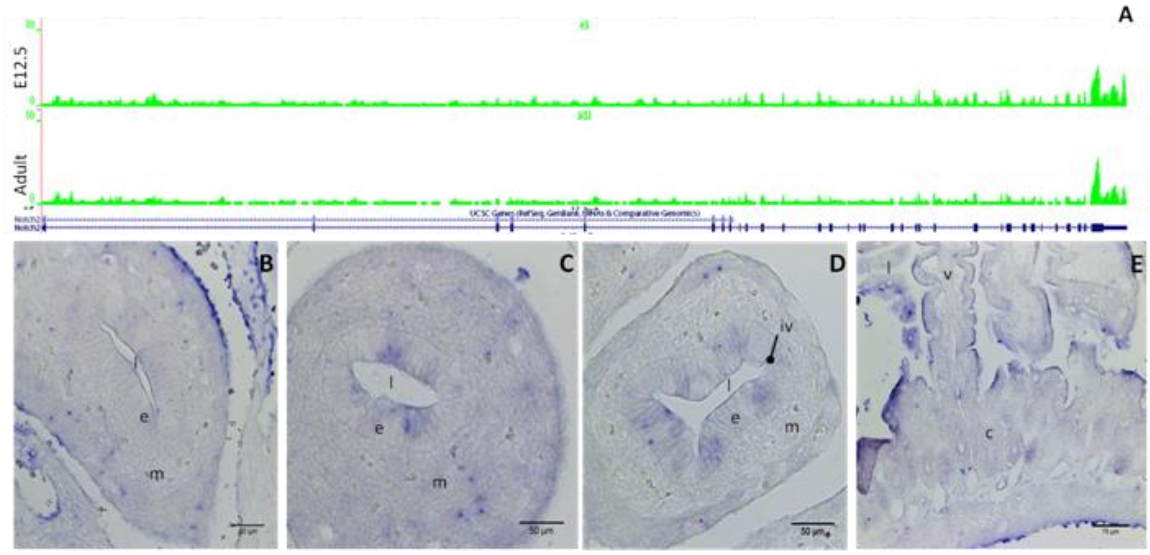


Fig 4. 14: Expression pattern of the gene Notch2 resulting from RNA-seq (A) and ISH (B to E). Data from RNA-seq (A) show the expression pattern between mouse e12.5 and adult ISC. Note the exon represented with the solid boxes at the bottom. ISH specific for the gene Notch2 in fixed sections of mouse embryos e13.5 (B), e14.5 (C), e15.5 (D) and adult mouse small intestine (E). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50µm to all sections except E, where it corresponds to 75µm.

Expression pattern of the gene Notch2 was performed by RNA-seq and ISH (Fig 4.14). Data from RNA-seq data (Fig 4.12 A) do not present major differences in the amount of transcripts between the ISCs populations from e12.5 and adult mice. Results from ISH on Notch2 (Fig 4.12 B to E) present changes in staining intensity in the intestinal endothelium. These changes are stronger at stages e14.5 and e15.5 (Fig 4.12 C and D, respectively), with slight changes being noticeable at e13.5 (Fig 4.12 B) and at the bottom of the crypts in section of the adult small intestine (Fig 4.12 E).

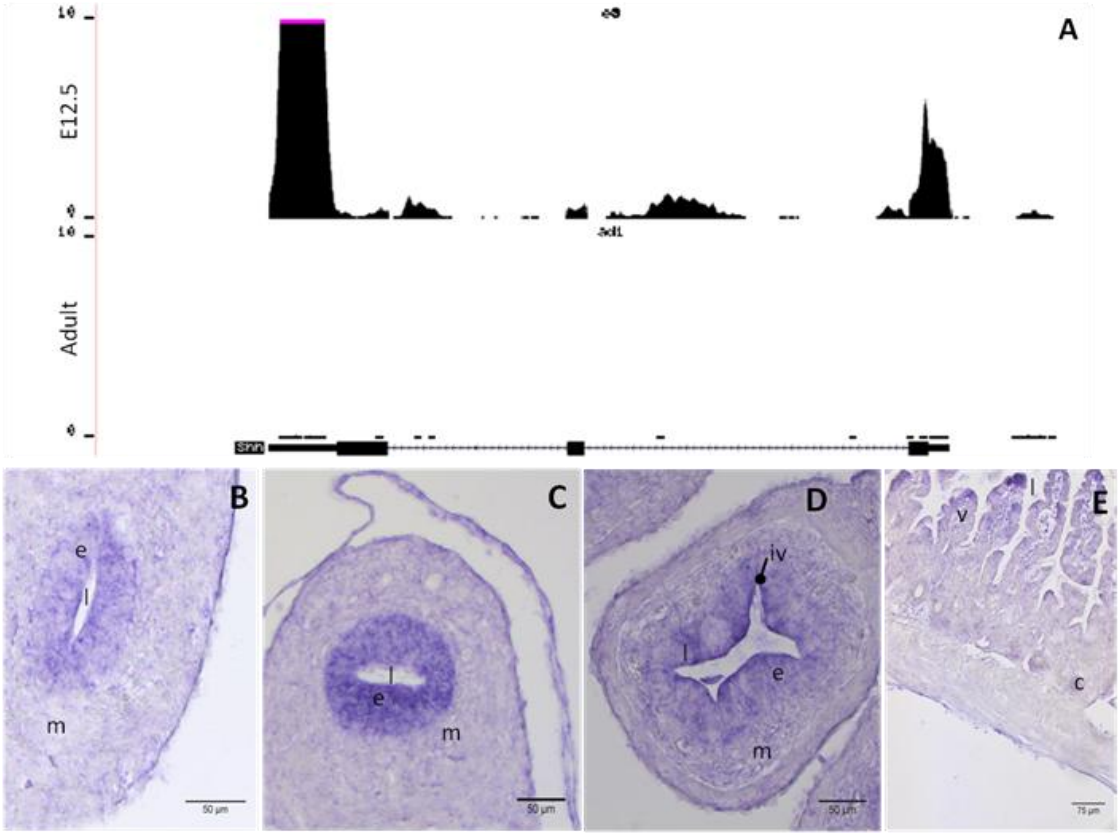


Fig 4. 15: Expression pattern of the gene Shh resulting from RNA-seq (A) and ISH (B to E). Data from RNA-seq – number of reads (A) show the expression pattern between mouse e12.5 and adult ISC. Note the exon represented with the solid box at the bottom. ISH specific for the gene Shh in fixed sections of mouse embryos e12.5 (B), e14.5 (C), e15.5 (D) and adult mouse small intestine (E). Note the different structures present in the sections throughout the developmental stages, l – lumen, e - endothelium, m – mesenchyme, iv – intravillus, v – villi, c – crypt. Scale bar corresponds to 50μm to all sections except E, where it corresponds to 75μm.

According to the RNA-seq data, the expression of Shh in the ISC of the mouse embryo, e12.5, and in the adult small intestine, is higher at early developmental stages (Fig 4.15 A). This is confirmed by the ISH analysis of this gene (Fig 4.15 B to E). In the images from the ISH it can be observed a stronger staining in the intestinal endothelium at the embryonic stages, e12.5, e14.5 and 15.5, analyzed (Fig 4.15 B to D). In the section from the adult small intestine (Fig 4.15 E) it is not possible to identify a region, in the crypts, where the staining is stronger, even though the background is slightly strong.

The probe for the gene Ngn3 was prepared and the ISH performed. Nevertheless, no specific signal was produced during the ISH of slides targeting this gene transcripts. For this reason, no results were presented.

4.3 Isolation of Intestinal cells

The isolation of the target cell populations, ISC and AE, required upstream steps prior to cell sorting. To successfully isolate the embryonic ISC, independently of the embryonic stage, the small intestine was dissected out and isolated from the embryos (Fig 4.16 A and B) . For the isolation of adult ISC and AE population, the small intestine was collected from male Lgr5-EGFP-ires-creERT2 and sectioned (Fig 4.16 C), prior to further processing.

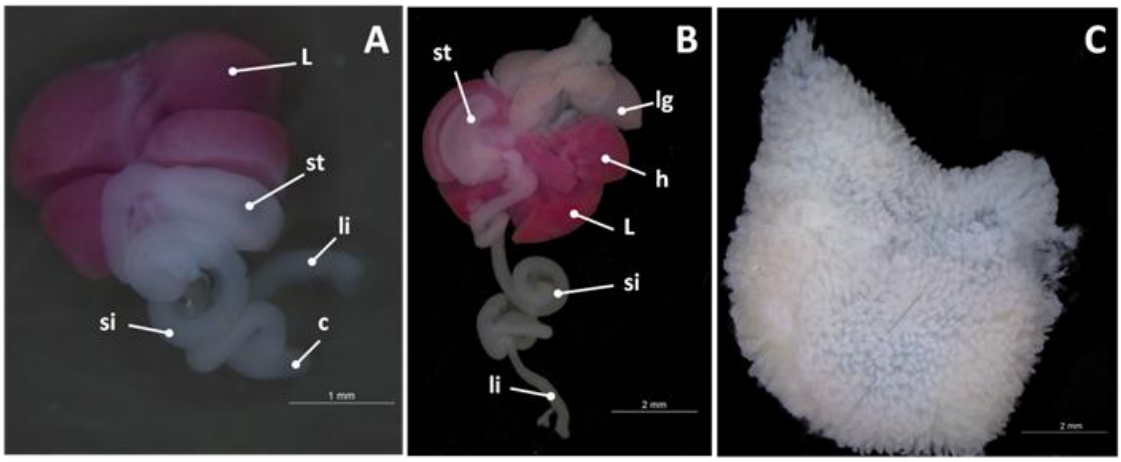


Fig 4. 16: Mouse embryos abdominal organs at developmental stages e12.5 (A) and e14.5 (B) and small section of adult small intestine (C). The abdominal organs collected included liver (L), stomach (st), small intestine (si), large intestine (li), lungs (lg), heart (h) and in the tissues collected from e12.5 (A) it's possible to identify the ceccum (c). Scale bars are 1 mm for insert A and 2 mm for inserts B and C.

4.3.1 Isolation of adult ISCs and AE

The isolation of adult ISCs and AE were carried out in male Lgr5-EGFP-ires-creERT2. This mouse strain has the characteristic of expressing EGFP in the cells expressing Lgr5 – ISCs.

In order to ensure the purity of the isolated cells, they were stained with EpCAM, CD31 and C45 antibodies. This antibody combination allowed the identification of dendritic and other immune cells are abundant in the cell suspentions. Cells that were co-stained by these three antobodies were rejected in order to obtain clean preparations of adult ISCs and AE . Cells that were GFP and EpCAM positive, with a high expression of this protein were collected as the ISCs of interest (Fig 4. 17). On the other hand, cells only stained with EpCAM corresponded to the targeted AE (Fig 4.18).

Cells were also stained with DAPI to descriminate between the living and dead cells.

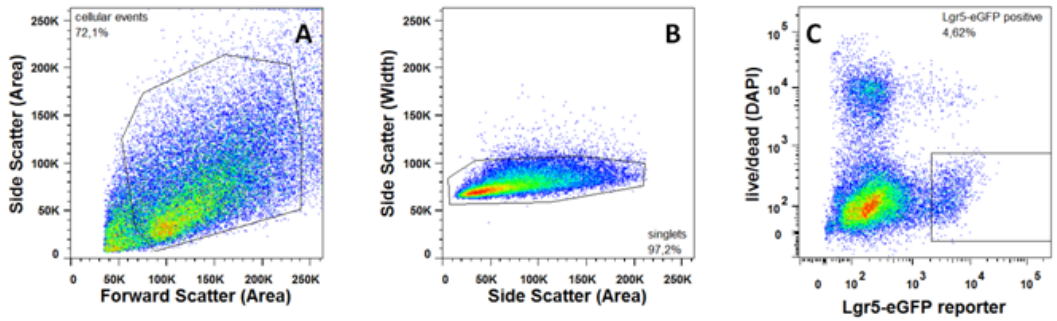


Fig 4. 17: Cell sorting (FACS) of adult ISCs from male Lgr5-EGFP-ires-creERT2 stained with EpCAM, CD31 and CD45, and DAPI. Cellular events (A), single cell events (B) and Lgr5-eGFP positive events (C) were isolated in order to prepare a purified cell suspention of the targeted adult ISCs.

From the cell suspension loaded into the FACS equipment, cells that were positive for all the antibodies were isolated. Out of this cell suspension, cells were isolated from cellular debris resulting from the preparation (Fig 4.17 A). This accounted for 72.1% of all events. The isolated material was further screened in order to discard events that were not single cell (Fig 4.17 B). From the all the cellular events, 97.2% were single events. Of these, the cells that expressed the Lgr5-eGFP reporter and were alive were isolated (Fig 4.17C), accounting for 4.62% of cells

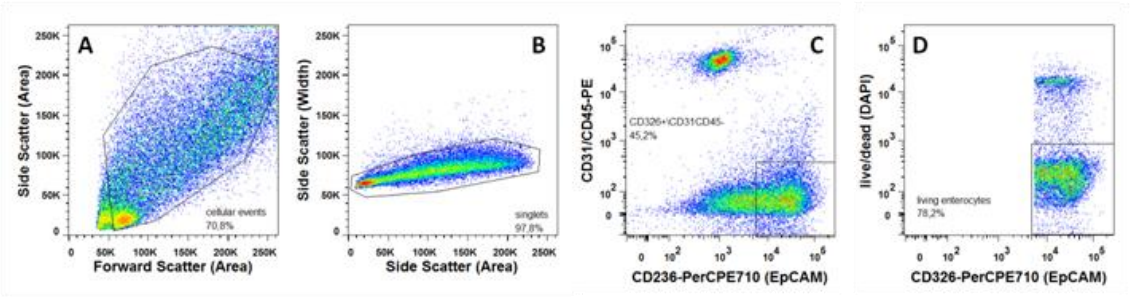


Fig 4. 18: Cell sorting (FACS) of adult enterocytes from male Lgr5-EGFP-ires-creERT2 stained with EpCAM, CD31 and CD45, and DAPI. Cellular events (A), single cell events (B) EpCAM, CD31 and CD45 stained cells (C). Living cells EpCAM positive and CD31/CD45 negative (D) were isolated in order to prepare a purified cell suspention of the targeted adult ISCs.

The enterocyte cell suspension was loaded into the FACS equipment and 70.8% were celular events (Fig 4.18 A) to be collected out of the cellular debries and other contaminants. Single cell events accounted for 97.8% (Fig 4.18 B) of the cellular events present in the cell suspension. Cells positive for EpCAM but negative either for CD31 and CD45 were 45.2% of all single cell events (Fig 4.18 C) and 78.2% of them were alive (Fig 4.18 D) and thus collected for further applications.

4.3.2 Isolation of embryonic ISCs

The small intestines that were dissected from the embryos and pooled – 8 to 10 guts from each litter. To ensure the purity of the isolated cells, these were stained with the antibodies EpCAM and CD31 for the stage e12.5 and with EpCAM, CD31 and CD45 for the stage e14.5.

The isolation of the ISCs from stage e12.5 was carried out by collecting the cells that were EpCAM positive and CD31 negative (Fig 4.19). At this stage, e12.5, the EpCAM will stain the well characterized pan-endodermal surface marker. In order to complement the staining and increase the purity of the cells collected, they were also stained with CD31. In respect to the isolation of the ISCs at stage e14.5, this was

performed in such way that cells co-stained with CD31, CD45 and EpCAM were dismissed. Only cells stained with EpCAM were collected for further use (Fig 4.20).

Staining with DAPI was also performed to identify living cells, with no DAPI staining from the dead, stained with DAPI in the nucleus.

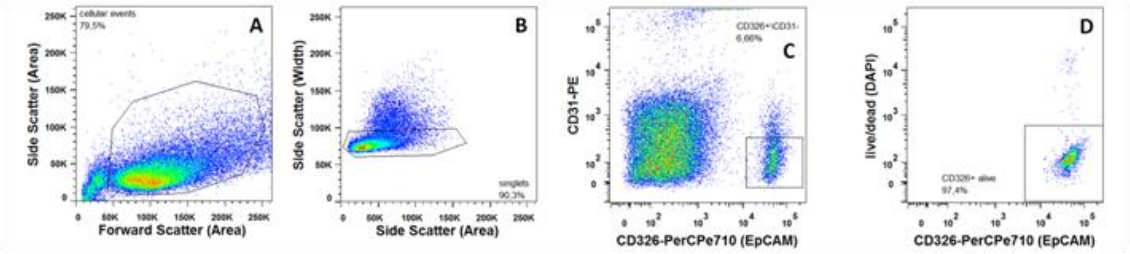


Fig 4. 19: Cell sorting (FACS) of embryonic endothelium cells from small intestine of stage e12.5, stained with EpCAM and CD31. Target cells were EpCAM+/CD31-. Cellular events (A), single cell events (B) EpCAM, and CD31 stained cells (C). Living cells EpCAM positive and CD31 negative (D) were isolated in order to prepare a purified cell suspention of the targeted embryonic ISCs.

From the FACS sorting, in the cell suspension loaded into the equipment, cells were isolated from the cellular debries produced during the preparation, wich accounted for 79.5% of all events (Fig 4.19 A). The suspension was further screened to discard the events that were not single cell collecting 90.3% of all the events (Fig 4.19 B). Out of these, cells EpCAM positive and CD31 negative accounted 6.66% (Fig 4.19 C), with 97.4% of them being alive (Fig 4.19 D).

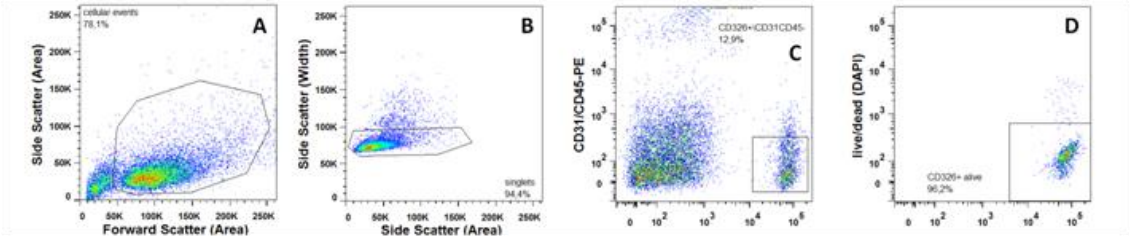


Fig 4. 20: Cell sorting (FACS) of embryonic endothelium cells from small intestine of stage e14.5, stained with EpCAM, CD31 and CD45. Target cells were EpCAM+/CD31CD45-. Cellular events (A), single cell events (B) EpCAM, CD31 and CD45 stained cells (C). Living cells EpCAM positive and CD31/CD45 negative (D) were isolated in order to prepare a purified cell suspention of the targeted embryonic ISCs.

The cell suspension loaded into the FACS equipment allowed the recovery of 78.1% of cellular events out of the cell debries and other contaminants (Fig 4.20 A). Single cell events accounted for 94.4% (Fig 4.20 B) of all cellular events in the suspension. Cells positive for EpCAM but negative either for CD31 and CD45were 12.9% of all single cell events (Fig 4.20 C) and 96.2% of them were alive (Fig 4.20 D) and thus collected for further applications.

4.4Bisulfite sequencing

In order to proceed for the bisulfite sequencing, cells that were sorted by FACS were processed. Genomic DNA was extracted and sonicated and sodium bisulfate convertion was carried out.

Simoultaneously, target genes were selected to be further analysed using information from the MDS-seq protocoled that was carried out. Genes to be analyzed were selected in accordance to their methylation profile.

Genes that were selected were Ndufaf3 (Fig 4.21), Id2 (Fig 4.22), Foxa1 (Fig 4.23), Nrp2 (Fig 4.24), Shh (Fig 4.25), Fzd2 (Fig 4.26), Meis1 (Fig 4.27), Grb10 (Fig 4.28), Mdk (Fig 4.29), Olfm4 (Fig 4.30), Elf3 (Fig 4.31), Ephb3 (Fig 4.32), Vdr (Fig 4.33), Oct4 (Fig 4.34), Lgr5 (Fig 4.35) and Efbn2 (Fig 4.36).

The following images are repesentative of each stage and method. Replicates were n=3 for the different stages at each protocol – RNA and MBD-seq.

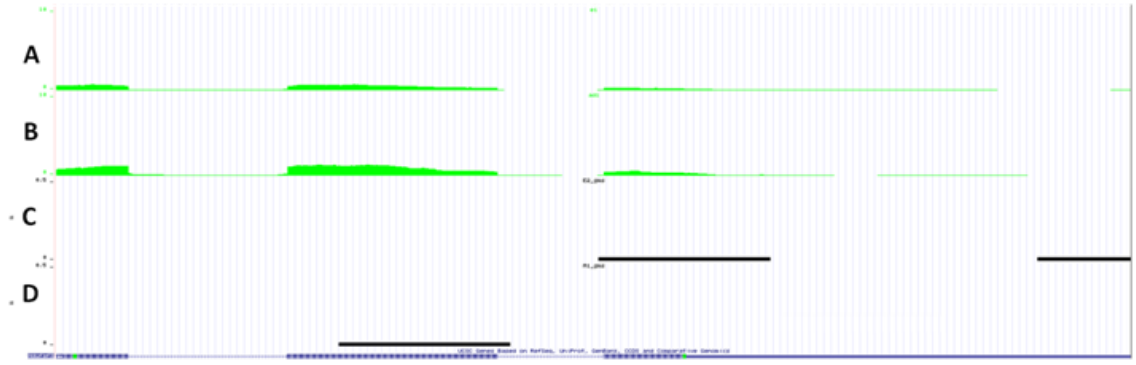


Fig 4. 21: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Ndufaf3 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exon represented with the solid blue boxes at the bottom.

Expression and methylation patterns of the gene Ndufaf3 (Fig 4.21), resulting from RNA-seq and MBD-seq. This data show a small increase of mRNA (Fig 4.21 A and B) in the adult ISC (Fig 4.21 B), when compared to the same cells from stage e12.5 (Fig 4.21 A). These changes were not observed in the methylation of the gene (Fig 4.21 C and D).

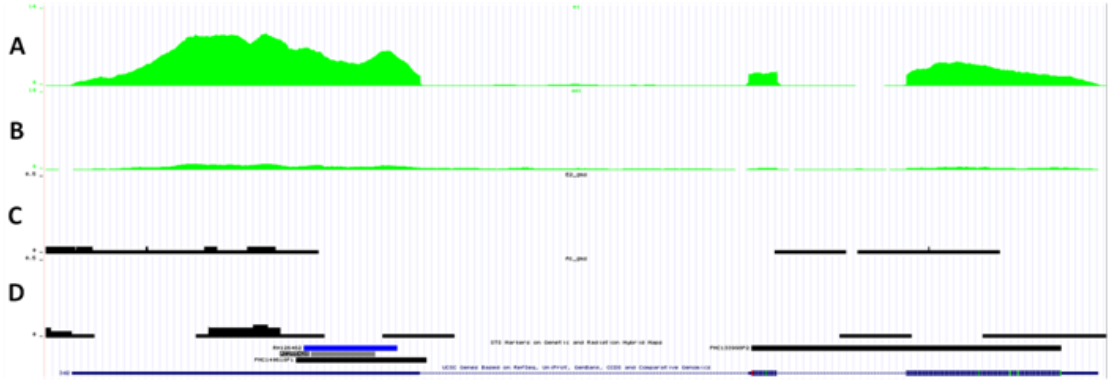


Fig 4. 22: Expression (lines A and B) and methylation (lines C and D) profiles throughout the gene Id2 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exon represented with the solid blue boxes at the bottom.

Analyzing the data from RNA-seq regarding the gene Id2 (Fig 4.22 A and B), it can be noticed that the expression of this gene is higher in the embryo, at stage e12.5 then it is in the adult ISC. Nevertheless, no changes in the methylation pattern (Fig 4.22 C and D) can be detected between the different developmental stages.

Fig 4. 23: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Foxa1 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exon represented with the solid box at the bottom.

Focusing in the RNA-seq and MBD-seq data for the gene Foxa1 (Fig 4.23) it can be noticed differences in the expression profile of this gene (Fig 4.23 A and B). Regarding this data, a higher expression of Foxa1 is observed from the ISC at stage e12.5 (Fig 4.23 A). On the other hand, the expression decreases in the analysis of the adult ISC (Fig 4.23 B). Concerning the methylation (Fig 4.23 C and D), data shows that there is no methylation in either stages.

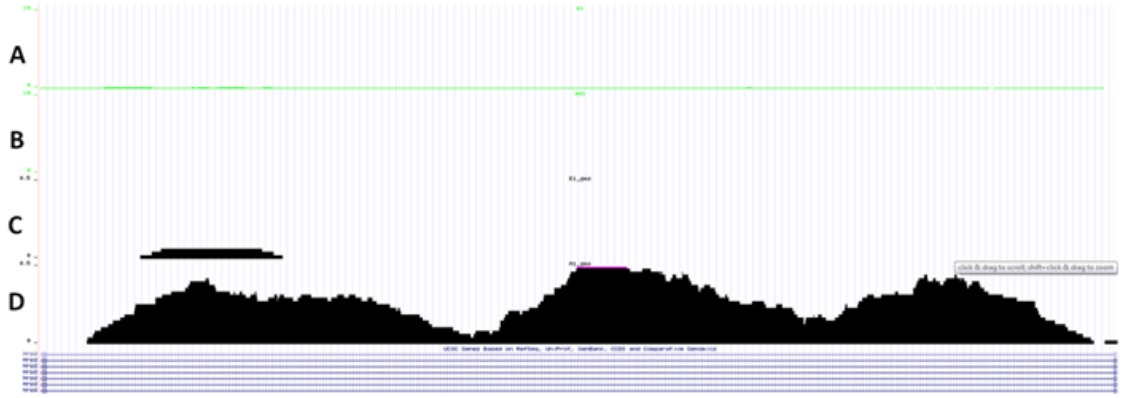


Fig 4. 24: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Nrp2 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D).

Data from RNA-seq and MBD-seq on Nrp2 (Fig 4.24) indicates that the expression of this gene in low for the stage e12.5 (Fig 4.24 A) and lower or inexistent in the adult ISC

(Fig 4.24 B), that may be due to the fact that the represented region corresponds to an intron. In respect to the methylation pattern (Fig 4.24 C and D), this is much higher in the adult ISC (Fig 4.24 D) and very low in at e12.5 (Fig 4.24 C).

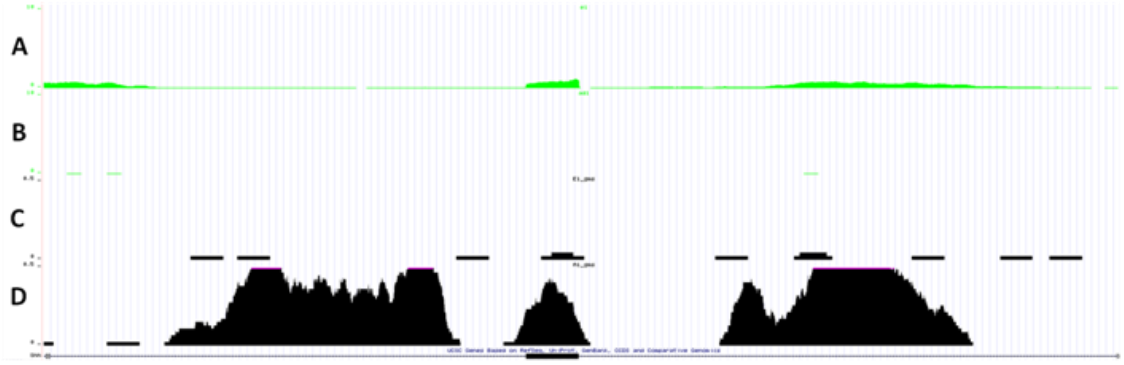


Fig 4. 25: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Shh in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exon represented with the solid box at the bottom.

Data from RNA-seq and MBD-seq for the gene Shh (Fig 4.25) gives a representation of the expression and methylation patterns. Regarding the expression pattern, transcripts are higher in the ISC at e12.5 (Fig 4.25 A) specially over the exon (bottom solid box, Fig 4.25) and inexistent in the adult ISCs (Fig 4.25 B). This gene present a high methylation in the adult ISCs (Fig 4.25 D) in the exon and also in the surrounding regions. In the e12.5 ISCs, methylation is very low (Fig 4.25 C).

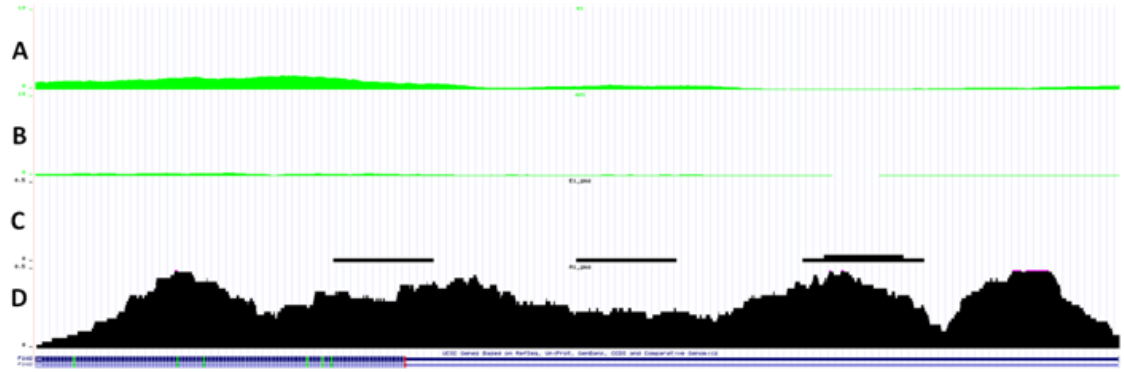


Fig 4. 26: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Fzd2 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Observing the RNA-seq and MBD-seq data regarding the gene Fzd2 (Fig 4.26) it can be noticed a higher expression of the gene at e12.5 (Fig 4.26 A) and a very low or inexistent methylation for the same stage (Fig 4.26 C). For the adult ISC, the opposite can be observed – low amount of transcripts (Fig 4.26 B) and highly methylated regions (Fig 4.26 D).

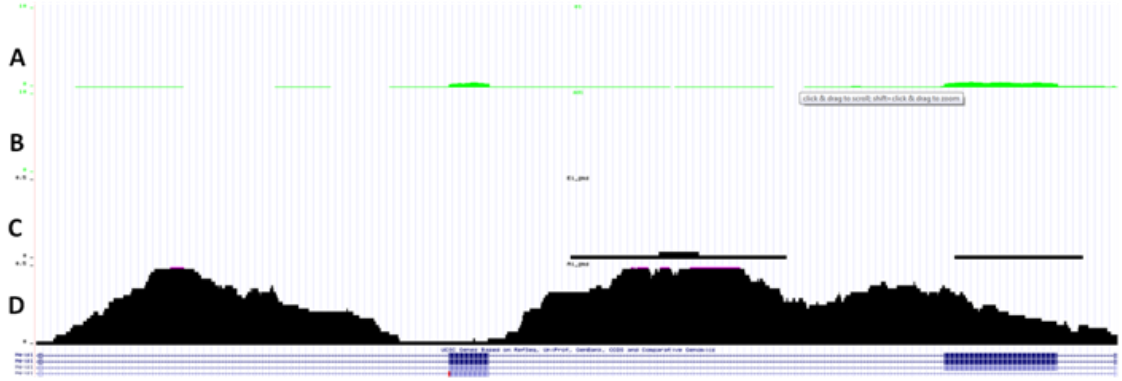


Fig 4. 27: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Meis1 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Focusing in the RNA-seq and MBD-seq data for the gene Meis1 (Fig 4.27) it can be noticed an expression at stage e12.5 (Fig 4.27 A) although low, that is not present in the adult ISCs (Fig 4.27 B). Regarding methylation, this is higher in the adult cells (Fig 4.27 D), where no expression was detected, while being very low in the e12.5. stage (Fig 4.27 C).

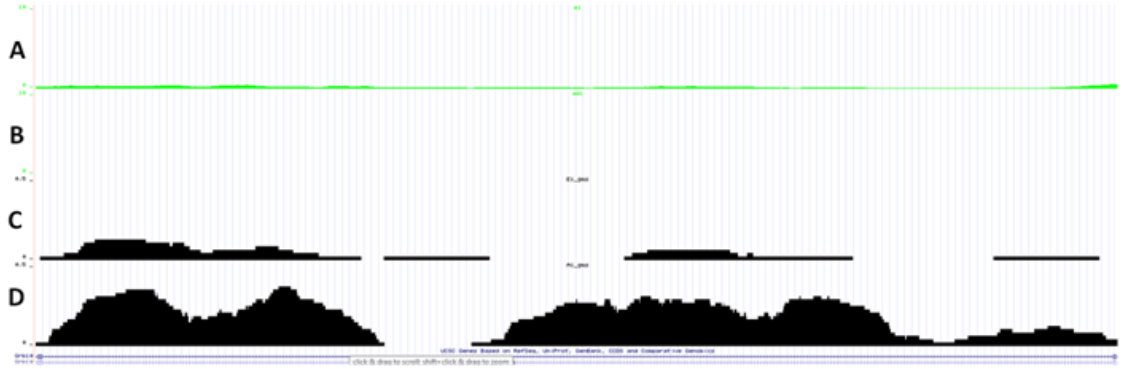


Fig 4. 28: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Grb10 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D).

Data from RNA and MBD-seq concerning Grb10 (Fig 4.28) show expression of the gene only on the embryonic stage, e12.5 (Fig 4.28 A), as no transcripts were detected in the adult ISC (Fig 4.28 B). Nevertheless, it is important to mention that the region of the gene being portrayed reefers to an intron. In respect to the methylation of the gene Grb10, it is possible to observe that this is higher in the adult ISC (Fig 4.28 D). Nonetheless, methylation was detected in the DNA from cells at e12.5 (Fig 4.28 C) although at low levels.

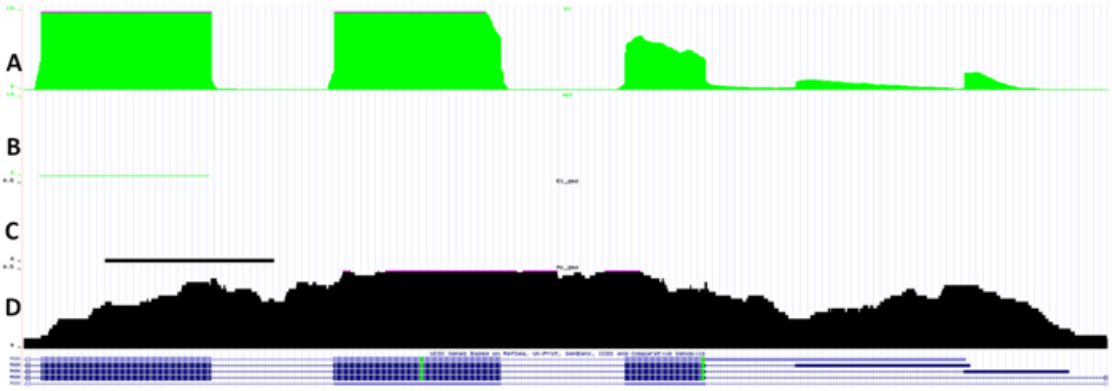


Fig 4. 29: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Mdk in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Expression and methylation pattern on the gene Mdk show differences between the two developmental stages analyzed (Fig 4.29). Regarding gene expression, RNA-seq data show that Grb10 is present in the e12.5 stage (Fig 4.29 A) and with a higher number of transcripts corresponding to the gene's exons. In the adult ISCs, no transcripts were detected (Fig 4.29 B). In respect to the MBD-data, no methylation was detected in the embryonic stage (Fig 4.29 C) while the gene was methylated in the adult ISCs in the introns and exons alike (Fig 4.29 D).

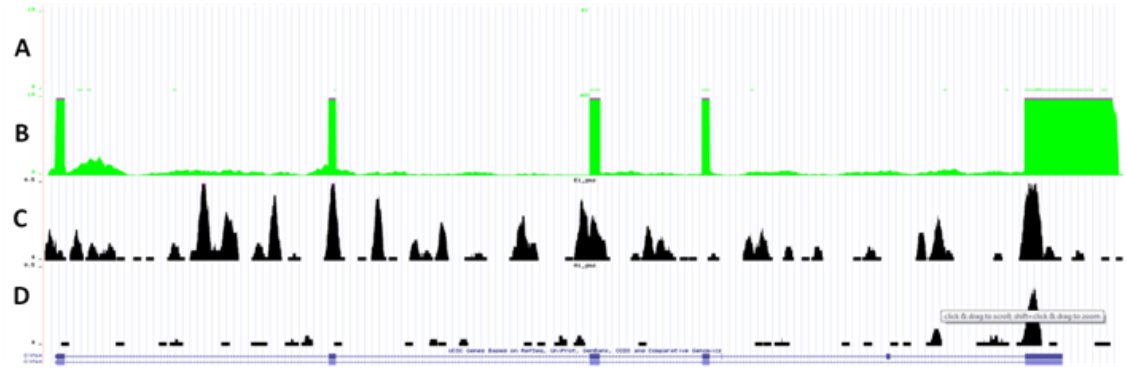


Fig 4. 30: Expression (lines A and B) and methylation (lines C and D) profiles throughout the gene Olfm4 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Analyzing the expression and methylation data for the gene Olfm4 (Fig 4.30) differences can be noted between the two developmental stages studied. In respect to the expression profile of Olfm4 (Fig 4.30 A and B), it can be noticed that this gene is expressed in the adult ISCs (Fig 4.30 B) while being inexistent in the embryonic ISCs at stage e12.5 (Fig 4.30 A). Regarding the methylation (Fig 4.30 C and D), the opposite can be observed. Methylation in the embryonic stage (Fig 4.30 C) and reduced in the adult cells (Fig 4.30 D).

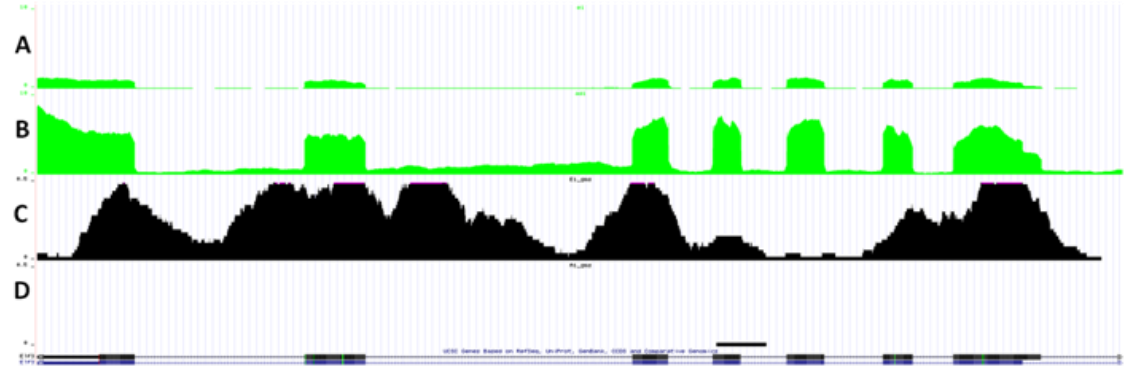


Fig 4. 31: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Elf3 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Focusing on the data for the gene Elf3 from RNA and MBD-seq (Fig 4.31) changes in expression and methylation can be observed between the stages studied. The gene Elf3

has a higher expression in the adult ISCs (Fig 4.31 B). Although this gene also appears to be expressed in the embryonic ISCs at e12.5 (Fig 4.31 A), transcript levels are low. On the other hand, DNA from e12.5 is highly methylated (Fig 4.31 C), while no methylation is detected in the adult ISCs (Fig 4.31 D).

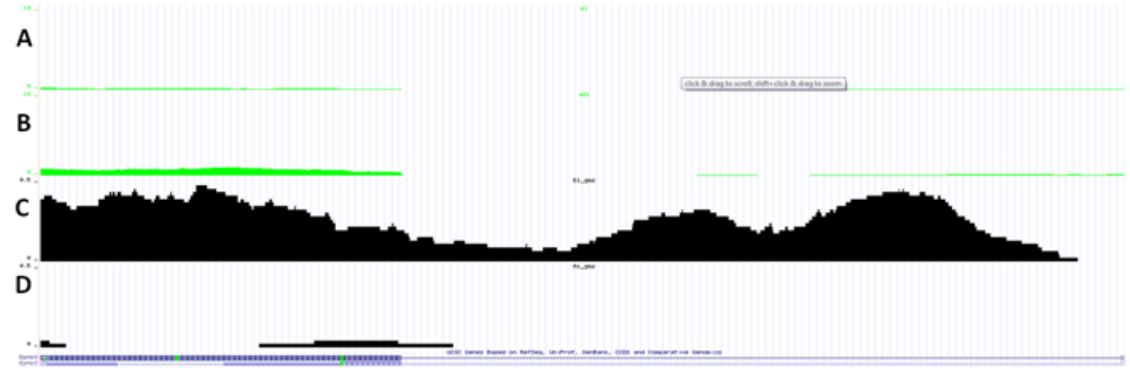


Fig 4. 32: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Ephb3 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

In respect to the gene Ephb3, RNA and MBD-seq data (Fig 4.32) show that there are differences between the two stages analyzed. Based in RNA-seq Ephb3 was showed to be expressed in the adult ISCs (Fig 4.32 B) and to be expressed at lower levels in stage e12.5 (Fig 4.32 A). MBD-seq data indicate a strong methylation of DNA in the embryonic ISCs (Fig 4.32 C) while being absent in the adult ISCs (Fig 4.32 D).

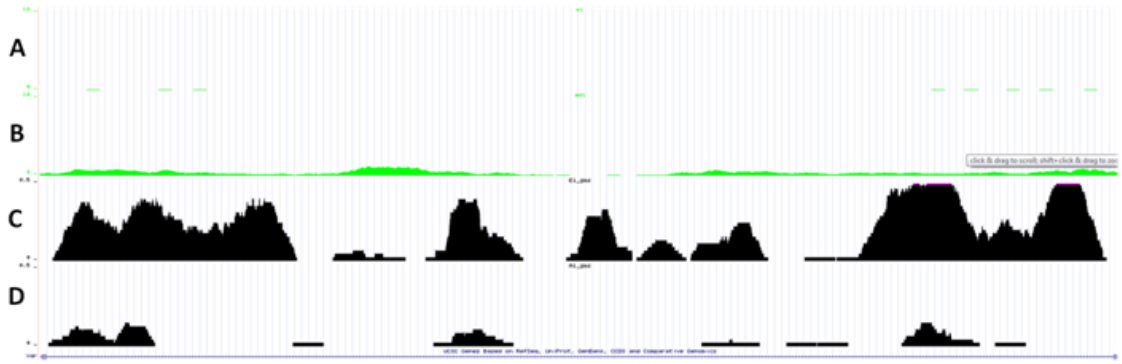


Fig 4. 33: Expression (lines A and B) and methylation (lines C and D) profiles of a region (an intron) of the gene Vdr in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D).

Analyzing the RNA-seq data related to the expression of Vdr (Fig 4.33 A and B) it can be noticed a higher amount of transcripts in the adult ISCs (Fig 4.33 B) while none were detected in the samples from e12.5 (Fig 4.33 A). Data regarding DNA methylation also show differences among the different stages (Fig 4.33 C and D). MBD-seq data show a higher methylation on the DNA from stage e12.5 (Fig 4.33 C). It also shows that some methylation in the adult ISCs have occurred (Fig 4.33 D), although lower than the detected for the embryonic ISCs.

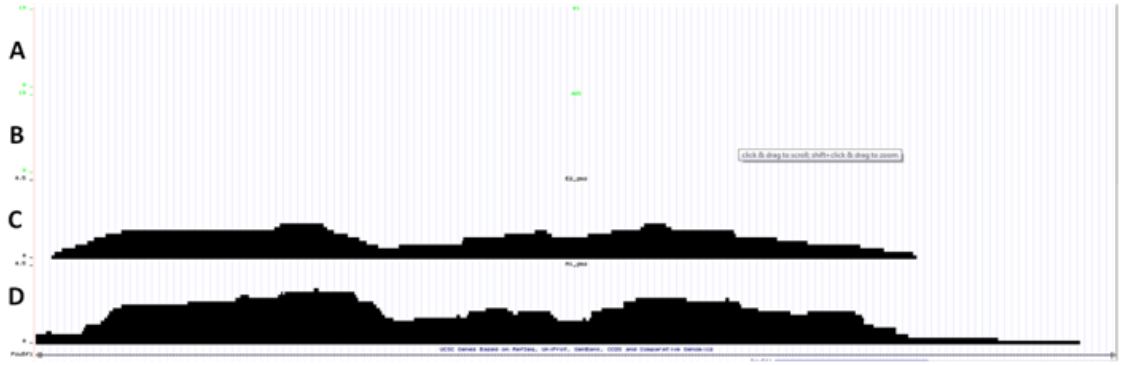


Fig 4. 34: Expression (lines A and B) and methylation (lines C and D) profiles of a region (an intron) of the gene Oct4 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D).

In respect to the gene Oct4, the RNA-seq protocol did not detect any transcripts present neither at stage e12.5 (Fig 4.34 A) neither in the adult ISCs (Fig 4.34 B). Nevertheless, the absence of transcripts in the region of the gene present in Fig 4.34 can be due to the fact that the region presented is an intron of the gene. Oct4 methylation is present in both stages (Fig 4.34 C and D).

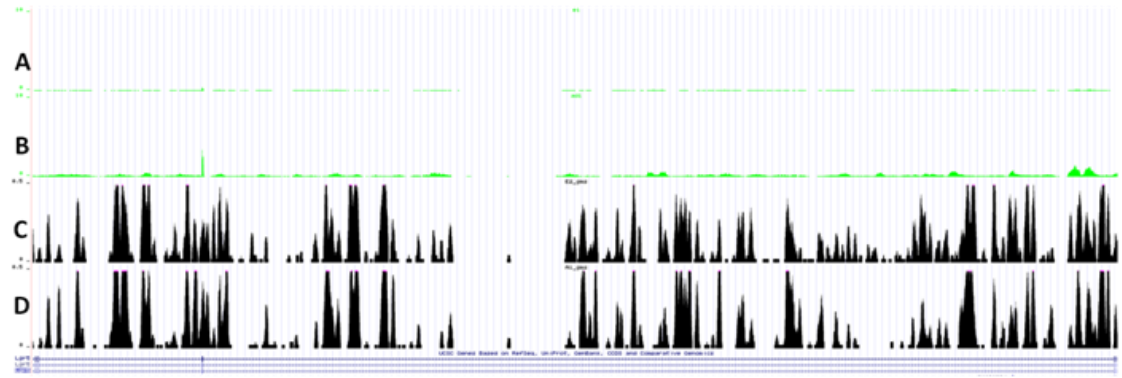


Fig 4. 35: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Lgr5 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Expression and methylation pattern on the gene Lgr5 show that while differences between the ISCs of the two developmental stages analyzed – e12.5 and adult (Fig 4.35) are present, the same is not noticed in the methylation. From the RNA-seq, it can be noticed a higher expression of the gene in the adult ISCs (Fig 4.35 B), specially in the regions corresponding to the exons. A very low amount of transcripts can be detected in the data from the e12.5 ISCs (Fig 4.35 A). In respect to the methylation of the gDNA coding for Lgr5, small differences could be noted by comparing the embryonic and adult ISCs (Fig 4.35 C and D).

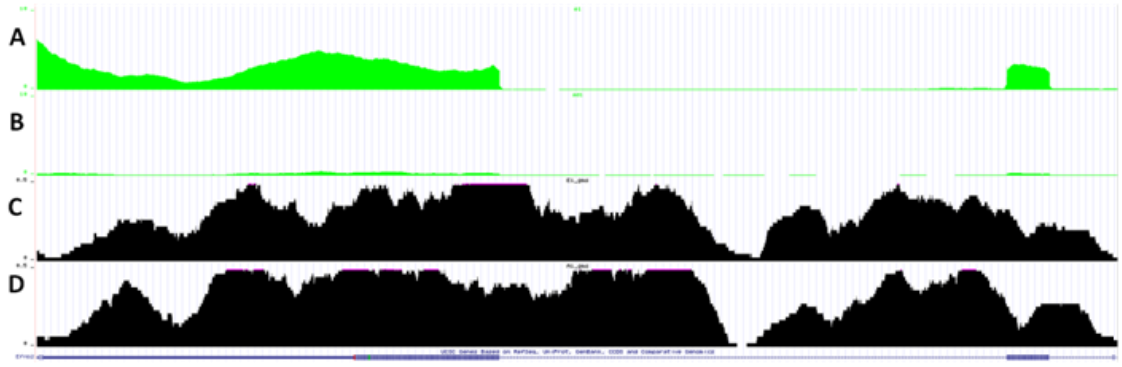


Fig 4. 36: Expression (lines A and B) and methylation (lines C and D) profiles of a region of the gene Efbn2 in ISC at developmental stage e12.5 (lines A and C) and adult (lines B and D). Note the exons as solid boxes at the bottom of the image.

Analysis of RNA and MBD-seq data regarding the gene Efbn2 (Fig 4.36) while indicating differences in transcript levels in the ISCs at e12.5 and adult – with expression at e12.5 (Fig 4.36 A) – show no major differences in the methylation pattern of DNA (Fig 4.36 C and D), being high in both stages.

Based on the MBD-seq data previously presented – Fig 4.21 to Fig 4.36, primers for the Bisulfite-seq were designed to amplify fragments that were unmethylated in both stages, methylated in adult and unmethylated in embryo, methylated in embryo and unmethylated in adult and methylated in both stages. This study was complemented with methylation analysis of DNA from ISC from stage e14.5 and from adult enterocytes.

Samples were treated with the sodium bisulfate followed by PCR in order to amplify the fragment of interest. Primer optimization was also taken as an opportunity to screen the primers that allowed fragment amplification. Products of the successful PCR reactions were loaded into an agarose gel (Fig 4.37).

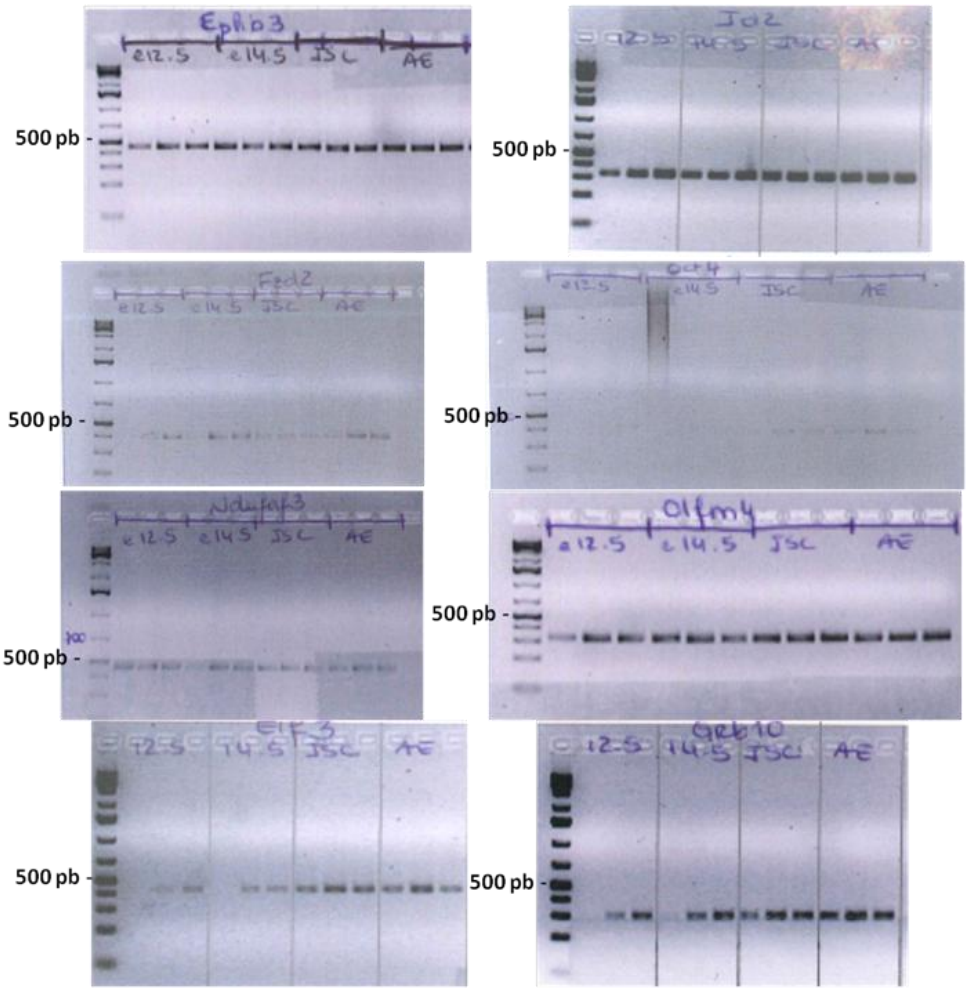


Fig 4. 37: Products of PCR reactions in a 1.5% agarose gel in 1X TAE stained with EtBr. PCR reaction was carried out for the amplification of several gene fragments – Ephb3, Id2, Fzd2, Oct4, Ndufaf3, Olfm4, Elf3 and Grb10 – after sodium bisulfate treatment of DNA. Triplicates of each developmental stage (e12.5, e14.5, adult ISC and AE) were prepared.

From the initial 16 genes selected, 8 were successfully amplified by PCR and the fragments obtained had sizes corresponding to the prediction when primers design (Table 3.10). PCR products were obtained for all samples, although band intensity could vary slightly.

The PCR fragments were to be quantified and pooled for each of the samples in order to proceed to sequencing. In parallel experiments that were being carried out in the group, it was detected an increase on available DNA in the samples being used for the bisulfate sequencing when compared with the samples used for the MBD-seq (data not shown). For this discrepancy, the fragments prepared for the bisulfate sequencing were not sequenced, as the results would not be comparable.

5. Discussion

Stem cells are able to self-renew and to specialize in multiple differentiated cell types[62]. These differentiated cells are what will give an organism the ability to respond to the diverse environmental, developmental and/or metabolic cues[69]. The process through which these highly specialized cells arise, begin at the stem cells and as they undergo crucial changes in gene expression, they differentiate and acquire their final identity.[62] However, it is important to keep in mind that although the transcripts present in a fully differentiated cell are highly specific, all cells in an organism share the same DNA between each other, even though they are located in different tissue or found at different developmental stages[69,76]. The mechanisms involved in the acquisition of such a tight expression program is the result of epigenetic modification to the DNA[62]. Results in the present thesis show that different developmental stages have different genetic programs being expressed and also, that, for the most cases, the modifications in the epigenetic marker DNA methylation are associated with the expression state of the gene the marker is associated with.

The main signaling pathways regulating the maintenance of the stem cells in the small intestine include Wnt/β-catenin, Notch, BMP and Hh. These signaling cascades have their own role in maintaining the cell niche. In the small intestine, Wnt/β-catenin is required for ISC maintenance and proliferation as well as for crypt compartmentalization and epithelium differentiation through a Wnt gradient that is established along the crypt-villus axis[4,5,27,34]. Notch, on the other hand, signals for the maintenance of the stem cell signal (suppressing cell fate decisions), cellular patterning and cell organization[4,36-38]. In the small intestine Notch pathway controls the differentiation of daughter cells into enterocytes while, together with Wnt, promotes ISC proliferation[3,39,40]. In respect to the BMP and Hh pathways, the first inhibits stem cell proliferation while the second is key for the normal embryonic development of the gut and its presence in the adult gut is related to stem cell maintenance, proliferation and differentiation[27,48,49].

Analysis of the RNA-seq data allowed the selection of set of genes potentially involved either in the development of the small intestine in embryos or the maintenance of the ISCs in the adult small intestine. An expression analysis was performed in this study by ISH and to detect the transcripts of Ascl2, Notch2, Hes1, Shh, Btg2, Cps1, Dusp1, Kcne3, Muc4 and Igfbp5.

The genes selected for ISH were among the many that, through RNA-seq analysis, presented different expression states in the stages analyzed. In fact, the adult ISCshowed transcription of several of the genes selected. The genes identified in the adult cells were Ascl2, Notch2, Hes1, Btg2, Cps1, Dusp1, Kcne3 and Muc4. It was interesting to note that from these genes only Ascl2, Notch2 and Hes1 were part of the major signaling pathways, with Ascl2 being a part of Wnt pathway and Notch2 and Hes1 being a part of Notch pathway[4,37,38,77]. In respect to the remaining genes, which were expected to be involved processes such as cell growth and/or proliferation, they were not directly associated with a pathway with the exception of Dusp1 which is involved in the MAPK pathway, thus showing that although Notch, Wnt and the other major signaling pathways are relevant for the ISC maintenance and differentiation, they are not the only pathways involved.

Signal from the ISH associated with the transcripts of Ascl2, Btg2, Cps1, and Muc4 was only detected in the adult ISCs (Fig 4.6, 4.7, 4.8 and 4.13, respectively), while transcripts were detected throughout all the embryonic stages analyzed (e12.5, e13.5, e14.4 and e15.5) and adult in the ISH targeting Dusp1, Kcne3, Notch2 and Hes1. This distribution of the transcripts can be associated with the functions of each of the genes. The first group, Ascl2, Btg2, Cps1 and Muc4, include a gene belonging to the Wnt/β-catenin pathway (Ascl2), genes implicated in cell cycle arrest (Btg2)[78], another gene who unique activity is ammonia metabolization (Cps1)[79,80] and Muc4, an epithelial mucin that can also regulate epithelial renewal and differentiation[81]. All of these genes (with the exception of Cps1) have a role in the maintenance of cells found in the adult stem cell niche. Nevertheless, in a previous work from van Beers *et al*(1998) Cps1 transcripts were detected in the crypts of adult rat small intestine[79] which is in accordance with observations from the present study. In the second set of genes are Notch2 and Hes1 components of the Notch pathway and Dusp1 that is involved in the control of genes promoting cell proliferation and differentiation[82,83]. The exception in

this group is Kcne3, a subunit of a potassium chanel[84] with no data found associating this gene with stem cell maintenance and/or differentiation.

Signal in the ISH targeting the genes Igfbp5 and Shh was not detected in adult ISCs (Fig 4.11 and 4.15) but instead in the embryonic stages. The results in this study regarding Shh are in accordance with the expected as it has been described that Shh is expressed throughout the embryonic epithelium, becoming progressively restricted into the proliferative regions of the small intestine[13,54]. In respect to the Igfbp5, this gene is also involved in the development of the small intestine by regulating the IGF and consequently cellular growth and differentiation[85,86]

With the ISH it was possible to validate RNA-seq data while allowing to better understanding of how these genes integrate the gene programs leading to gut development (in embryo) and the maintenance of the small intestine. It was not surprising the fact that the majority of the genes analyzed had a role in cell maintenance, proliferation or differentiation, despite belonging or not to one of the major signaling pathways. What was interesting was to have detected, in this study, the genes Cps1 and Kcne3 being detected in these cell populations when no information so far has related these genes with the processes that occur in this cells.

RNA-seq data and ISH results have showed that gene expression changes through the development of the small intestine. Genes that are transcribed at one moment may not be in the other. One of such cases is the gene programs in the embryonic development.

During embryogenesis cells begin as totipotent cells with the potential do differentiate in any type of cell. Progressively cells undergo changes in gene expression, which, in turn, causes for several factors to be progressively silenced[62]. This silencing result in a more specific cell type. Such cell type can be an adult stem cell that although able to divide and multiply is already committed to a lineage, or a fully differentiated cell – not able to divide or multiply, but fully adapted to the tissue and able to respond to different environmental, developmental and metabolic cues[62,69]. The selective expression at this moment, with the silencing of some factors and activation of others, is not the result of changes in DNA sequence but are a consequence of heritable epigenetic rearrangements

of DNA[62,64]. As epigenetic reinforce the decisions regarding cell-fate that were taken, they also create a barrier against a setback towards preceding cell states[87].

Epigenetic changes include DNA methylation, chromatin remodeling and histone modifications, to name a few[63-65,87,88]. From the different epigenetic marks, the work developed in this thesis focused in DNA methylation.

Results from MBD-seq allowed to collect data regarding the epigenetic status of DNA, focused on DNA methylation, in the ISCs isolated from embryos at stage e12.5 and adult small intestine. Some genes were selected for further scrutiny by bisulfite sequencing. Those genes were Shh, Grb10, Foxa1, Id2, Ndufaf3, Nrp2, Fzd2, Meis1, Mdk, Efbn2, Lgr5, Vdr, Olfm4, Elf3, Ephb3 and Oct4.

As expected since the cell populations focused in this study were stem cells from the small intestine. Selection of these genes were based in 1) the stage they were expressed – e12.5 and/or adult – using the data from RNA-seq and 2) the methylation status of each gene in the different stages, which was determined by MBD-seq data analysis (Fig 4.21 to Fig 4.36).

In general, the genes expressed in the embryonic (e12.5) ISC – Shh, Grb10, Foxa1, Id2, Ndufaf3, Nrp, Fzd2, Meis1, Mdk and Efnb2 – were unmethylated at this stage. The exception was Efbn2 that although methylated and expressed at this stage. It is also to be noted that there were genes (Efbn2) that maintained their unmethylated status although inactive in the adult ISCs. In respect to the genes expressed in the adult ISC – Lgr5, Vdr, Olfm4, Elf3 and Ephb3 – there was also a gene that although methylated, was expressed at this stage, Lgr5. Also among the genes expressed in the adult ISCs, there were some, Ndufaf3, Id2 and Foxa1,that were unmethylated at e12.5.

Through this work, it was possible to verify that DNA methylation has an important role in the regulation of gene expression in ISCs – the majority of genes were methylated when they were not expressed. This results are in accordance with what has been described since methylation is the more recurrent epigenetic modification in eukaryotes[88]. As this marks are present in gDNA, they inhibit transcription initiation by blocking the bind of transcription factors to DNA or by recruiting methyl-binding proteins (MBPs) changing the chromatin environment[88,89]. As the major changes in this patterns occur at periods of genome reprogramming – primordial germ cells and

preimplantation embryos – when marks are erased and reset[76], the results from e12.5 confirm the presence of methylation marks controlling cell differentiation towards a fully functioning organism. However, modifications on methylation detected between the ISCs at e12.5 and in adult is also in accordance with what has been described, that states that methylation can change as a response to genetic variation and environmental cues, among other factors[90].

For the majority of the genes analyzed in this study, expression changed between the different developmental stages. It was possible to establish a relation to their methylation pattern – genes that were being transcribed weren't methylated. However there were genes with a methylation profile that did not changealthough their transcription did. In this cases, it can be hypothesized that other epigenetic marker(s) is(are) involved in the regulation of gene transcription. One of the possible epigenetic markers which may be involved in the regulation of these genes is histone modifications. The role of histone modifications in gene transcription is a result of modifications that occur on specific residues such as methylations, phosphorylations, acetylations, sumoylations and ubiquitinations[69,91]. These histone modifications will have an impact in the nucleosome, and consequently alters chromatin condition, thus influencing transcriptional status of genes[69,92]. One of these cases has been described for the genes found in highly packed areas of chromatin. In such cases, the genes and their respective promoters are inaccessible and are not able to recruit the RNA polymerase remaining silenced[69]. The opposite has been reported in embryonic stem cells where chromatin is "open" due to a global high level of histone acetylation which maintains cells undifferentiated and pluripotent state[91,93].

5.1 Future perspectives:

Data presented in this study shows that different genes are required for stem cell from small intestine to maintain themselves or to differentiate. ISH also give strong indications that changes occur in a dynamic manner throughout development until maturity is reached. Analysis of epigenetic markers on a few selected genes, has provided data that relate gene transcription with DNA methylation where increase in

cytosine methylation was associated with a state of active transcription. From the 16 genes analyzed some did not present this relation.

Approaches that would enrich this study include:

- RNA-seq of ISC from intermediate embryonic stages, such as e14.5, and also from newborn mice in order to understand the moments when the genetic programs change; RNA-seq on adult enterocytes and transit amplifying cell may allow the comprehension of the signaling pathways involved in cell specification;

- MBD-seq on ISCs from e14.5, newborn, adult enterocytes and transit amplifying cells aiming to explain how this epigenetic markers change during development and cell specification;

- ChIP-seq on ISCs from embryos (e12.5 and e14.5) and adult small intestine as well as in adult enterocytes and transit amplifying cells. This, together with the MBD-seq data will enrich the information of the epigenetic control associated with gene transcription.

6. Conclusion

This study, focused in the intestinal stem cells, show that maintenance of stem cells in their niche is associated with specific gene programs. The signaling pathways that are active in these cells allow for the maintenance of stem cell population and give the cues for cell differentiation into tissue specific cells. However, these pathways are not present through all the developmental stages. Results present in this report indicate that during embryonic development gene programs are activated or deactivated.

In order to understand how genes were “turned on or off” without changes in DNA sequence, MBD-seq data showed that gene transcription is associated with changes in epigenetic markers. In this study DNA methylation profile was analyzed allowing the establishment of a relation according to which, the majority of the genes being transcribed are unmethylated, changing to a methylated state when the genes are deactivated. However, it was also possible to hypothesize that DNA methylation was not the only epigenetic marker present in these cells.

In face of these results it can be concluded that intestinal stem cells are maintained through several well known and specific pathways present from early developmental stages that modify as the genetic program narrows and also that this genetic program is under the control of epigenetic markers of great importance for ISC specification.

7. Bibliography

1. van der Flier, L.G. & Clevers, H. Stem cells, self-renewal, and differentiation in the intestinal epithelium. *Annu Rev Physiol* **71**, 241-260 (2009).
2. Barker, N., Bartfeld, S. & Clevers, H. Tissue-resident adult stem cell populations of rapidly self-renewing organs. *Cell Stem Cell* **7**, 656-670 (2010).
3. Umar, S. Intestinal stem cells. *Curr Gastroenterol Rep* **12**, 340-348 (2010).
4. Spence, J.R., Lauf, R. & Shroyer, N.F. Vertebrate intestinal endoderm development. *Dev Dyn* **240**, 501-520 (2011).
5. Barker, N., van de Wetering, M. & Clevers, H. The intestinal stem cell. *Genes Dev* **22**, 1856-1864 (2008).
6. Vereecke, L., Beyaert, R. & van Loo, G. Enterocyte death and intestinal barrier maintenance in homeostasis and disease. *Trends Mol Med* **17**, 584-593 (2011).
7. Barker, N., van Oudenaarden, A. & Clevers, H. Identifying the stem cell of the intestinal crypt: strategies and pitfalls. *Cell Stem Cell* **11**, 452-460 (2012).
8. Vanuytsel, T., Senger, S., Fasano, A. & Shea-Donohue, T. Major signaling pathways in intestinal stem cells. *Biochim Biophys Acta* (2012).
9. de Santa Barbara, P., van den Brink, G.R. & Roberts, D.J. Development and differentiation of the intestinal epithelium. *Cell Mol Life Sci* **60**, 1322-1332 (2003).
10. Yen, T.H. & Wright, N.A. The gastrointestinal tract stem cell niche. *Stem Cell Rev* **2**, 203-212 (2006).
11. Howell, J.C. & Wells, J.M. Generating intestinal tissue from stem cells: potential for research and therapy. *Regen Med* **6**, 743-755 (2011).
12. Noah, T.K., Donahue, B. & Shroyer, N.F. Intestinal development and differentiation. *Exp Cell Res* **317**, 2702-2710 (2011).
13. Crosnier, C., Stamataki, D. & Lewis, J. Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat Rev Genet* **7**, 349-359 (2006).
14. Takashima, S. & Hartenstein, V. Genetic control of intestinal stem cell specification and development: a comparative view. *Stem Cell Rev* **8**, 597-608 (2012).
15. Shaker, A. & Rubin, D.C. Intestinal stem cells and epithelial-mesenchymal interactions in the crypt and stem cell niche. *Transl Res* **156**, 180-187 (2010).
16. Fukuda, K. & Kikuchi, Y. Endoderm development in vertebrates: fate mapping, induction and regional specification. *Dev Growth Differ* **47**, 343-355 (2005).
17. Hauck, A.L*., et al.* Twists and turns in the development and maintenance of the mammalian small intestine epithelium. *Birth Defects Res C Embryo Today* **75**, 58-71 (2005).
18. Lewis, S.L. & Tam, P.P. Definitive endoderm of the mouse embryo: formation, cell fates, and morphogenetic function. *Dev Dyn* **235**, 2315-2329 (2006).
19. Roobrouck, V.D., Ulloa-Montoya, F. & Verfaillie, C.M. Self-renewal and differentiation capacity of young and aged stem cells. *Exp Cell Res* **314**, 1937-1944 (2008).
20. Ema, H. & Suda, T. Two anatomically distinct niches regulate stem cell activity. *Blood* **120**, 2174-2181 (2012).
21. Jaenisch, R. & Young, R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567-582 (2008).
22. Heath, J.K. Transcriptional networks and signaling pathways that govern vertebrate intestinal development. *Curr Top Dev Biol* **90**, 159-192 (2010).
23. Takashima, S., Gold, D. & Hartenstein, V. Stem cells and lineages of the intestine: a developmental and evolutionary perspective. *Dev Genes Evol* **223**, 85-102 (2013).

24. Ng, A. & Barker, N. Stem cells in epithelial renewal and cancer of the intestine. *European Gastroenterology & Histopathology Review* **7**, 154-159 (2011).
25. Lin, S.A. & Barker, N. Gastrointestinal stem cells in self-renewal and cancer. *J Gastroenterol* **46**, 1039-1055 (2011).
26. Rizk, P. & Barker, N. Gut stem cells in tissue renewal and disease: methods, markers, and myths. *Wiley Interdiscip Rev Syst Biol Med* **4**, 475-496 (2012).
27. Yeung, T.M., Chia, L.A., Kosinski, C.M. & Kuo, C.J. Regulation of self-renewal and differentiation by the intestinal stem cell niche. *Cell Mol Life Sci* **68**, 2513-2523 (2011).
28. Pinchuk, I.V., Mifflin, R.C., Saada, J.I. & Powell, D.W. Intestinal mesenchymal cells. *Curr Gastroenterol Rep* **12**, 310-318 (2010).
29. Roeder, I., Loeffler, M. & Glauche, I. Towards a quantitative understanding of stem cell-niche interaction: experiments, models, and technologies. *Blood Cells Mol Dis* **46**, 308-317 (2011).
30. Jin, Y.R. & Yoon, J.K. The R-spondin family of proteins: emerging regulators of WNT signaling. *Int J Biochem Cell Biol* **44**, 2278-2287 (2012).
31. Baarsma, H.A., Konigshoff, M. & Gosens, R. The WNT signaling pathway from ligand secretion to gene transcription: Molecular mechanisms and pharmacological targets. *Pharmacol Ther* (2013).
32. MacDonald, B.T., Tamai, K. & He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev Cell* **17**, 9-26 (2009).
33. Clevers, H. & Nusse, R. Wnt/beta-catenin signaling and disease. *Cell* **149**, 1192-1205 (2012).
34. Scoville, D.H., Sato, T., He, X.C. & Li, L. Current view: intestinal stem cells and signaling. *Gastroenterology* **134**, 849-864 (2008).
35. Leushacke, M. & Barker, N. Lgr5 and Lgr6 as markers to study adult stem cell roles in self-renewal and cancer. *Oncogene* **31**, 3009-3022 (2012).
36. Vooijs, M., Liu, Z. & Kopan, R. Notch: architect, landscaper, and guardian of the intestine. *Gastroenterology* **141**, 448-459 (2011).
37. Fortini, M.E. Notch signaling: the core pathway and its posttranslational regulation. *Dev Cell* **16**, 633-647 (2009).
38. Katoh, M. Notch signaling in gastrointestinal tract (review). *Int J Oncol* **30**, 247-251 (2007).
39. Perdigoto, C.N. & Bardin, A.J. Sending the right signal: Notch and stem cells. *Biochim Biophys Acta* (2012).
40. Madison, B.B. & Nakagawa, H. Delta force in intestinal crypts. *Gastroenterology* **140**, 1135-1139 (2011).
41. Kopan, R. & Ilagan, M.X. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* **137**, 216-233 (2009).
42. Schweisguth, F. Regulation of notch signaling activity. *Curr Biol* **14**, R129-138 (2004).
43. Bragdon, B*., et al.* Bone morphogenetic proteins: a critical review. *Cell Signal* **23**, 609-620 (2011).
44. Singh, A. & Morris, R.J. The Yin and Yang of bone morphogenetic proteins in cancer. *Cytokine Growth Factor Rev* **21**, 299-313 (2010).
45. Nohe, A., Keating, E., Knaus, P. & Petersen, N.O. Signal transduction of bone morphogenetic protein receptors. *Cell Signal* **16**, 291-299 (2004).
46. Zhang, J. & Li, L. BMP signaling and stem cell regulation. *Dev Biol* **284**, 1-11 (2005).
47. Li, Z. & Chen, Y.G. Functions of BMP signaling in embryonic stem cell fate determination. *Exp Cell Res* **319**, 113-119 (2013).
48. Saqui-Salces, M. & Merchant, J.L. Hedgehog signaling and gastrointestinal cancer. *Biochim Biophys Acta* **1803**, 786-795 (2010).
49. Varjosalo, M. & Taipale, J. Hedgehog signaling. *J Cell Sci* **120**, 3-6 (2007).

50. Kim, B.M. & Choi, M.Y. New insights into the role of Hedgehog signaling in gastrointestinal development and cancer. *Gastroenterology* **137**, 422-424 (2009).
51. de Sauvage, F. Hedgehog Signaling in Development and Disease. in *Handbook of Cell Signaling*, Vol. 3 (eds. Bradshaw, R.A. & Dennis, E.A.) 1879-1884 (Elsevier Inc., 2010).
52. Ingham, P.W., Nakano, Y. & Seger, C. Mechanisms and functions of Hedgehog signalling across the metazoa. *Nat Rev Genet* **12**, 393-406 (2011).
53. Katoh, Y. & Katoh, M. Hedgehog signaling pathway and gastrointestinal stem cell signaling network (review). *Int J Mol Med* **18**, 1019-1023 (2006).
54. Lees, C., Howie, S., Sartor, R.B. & Satsangi, J. The hedgehog signalling pathway in the gastrointestinal tract: implications for development, homeostasis, and disease. *Gastroenterology* **129**, 1696-1710 (2005).
55. Himanen, J.P. & Nikolov, D.B. Eph receptors and ephrins. *Int J Biochem Cell Biol* **35**, 130-134 (2003).
56. Arvanitis, D. & Davy, A. Eph/ephrin signaling: networks. *Genes Dev* **22**, 416-429 (2008).
57. Himanen, J.P. & Nikolov, D.B. Eph signaling: a structural view. *Trends Neurosci* **26**, 46-51 (2003).
58. Pasquale, E.B. Eph-ephrin bidirectional signaling in physiology and disease. *Cell* **133**, 38-52 (2008).
59. Arvanitis, D.N. & Davy, A. Regulation and misregulation of Eph/ephrin expression. *Cell Adh Migr* **6**, 131-137 (2012).
60. Edwards, C.M. & Mundy, G.R. Eph receptors and ephrin signaling pathways: a role in bone homeostasis. *Int J Med Sci* **5**, 263-272 (2008).
61. Mohn, F. & Schubeler, D. Genetics and epigenetics: stability and plasticity during cellular differentiation. *Trends Genet* **25**, 129-136 (2009).
62. Vincent, A. & Van Seuningen, I. Epigenetics, stem cells and epithelial cell fate. *Differentiation* **78**, 99-107 (2009).
63. Zhao, X., Ruan, Y. & Wei, C.L. Tackling the epigenome in the pluripotent stem cells. *J Genet Genomics* **35**, 403-412 (2008).
64. Meissner, A. Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* **28**, 1079-1088 (2010).
65. Bibikova, M., Laurent, L.C., Ren, B., Loring, J.F. & Fan, J.B. Unraveling epigenetic regulation in embryonic stem cells. *Cell Stem Cell* **2**, 123-134 (2008).
66. Tammen, S.A., Friso, S. & Choi, S.W. Epigenetics: the link between nature and nurture. *Mol Aspects Med* **34**, 753-764 (2013).
67. Collas, P. Epigenetic states in stem cells. *Biochim Biophys Acta* **1790**, 900-905 (2009).
68. Altun, G., Laurent, L.C. & Loring, J.F. Epigenetic remodeling and stem cells. *Drug Discov Today* **5**, 4 (2010).
69. Barrero, M.J., Boue, S. & Izpisua Belmonte, J.C. Epigenetic mechanisms that regulate cell identity. *Cell Stem Cell* **7**, 565-570 (2010).
70. Speel, E.J., Hopman, A.H. & Komminoth, P. Tyramide signal amplification for DNA and mRNA in situ hybridization. *Methods Mol Biol* **326**, 33-60 (2006).
71. Vize, P.D., McCoy, K.E. & Zhou, X. Multichannel wholemount fluorescent and fluorescent/chromogenic in situ hybridization in Xenopus embryos. *Nat Protoc* **4**, 975-983 (2009).
72. McNicol, A.M. & Farquharson, M.A. In situ hybridization and its diagnostic applications in pathology. *J Pathol* **182**, 250-261 (1997).
73. Glick, B.R. & Pasternak, J.J. *Molecular biotechnology: principles and applications of recombinant DNA*, (ASM Press, 2003).
74. Lodish, H.*, et al. Molecular cell biology*, (W. H. Freeman, New York, 2000).
75. Sun, W. & Hu, Y. eQTL Mapping Using RNA-seq Data. *Stat Biosci* **5**, 198-219 (2013).
76. Senner, C.E. The role of DNA methylation in mammalian development. *Reprod Biomed Online* **22**, 529-535 (2011).

77. Jarriault, S.*, et al.* Signalling downstream of activated mammalian Notch. *Nature* **377**, 355-358 (1995).
78. Guardavaccaro, D.*, et al.* Arrest of G(1)-S progression by the p53-inducible gene PC3 is Rb dependent and relies on the inhibition of cyclin D1 transcription. *Mol Cell Biol* **20**, 1797-1815 (2000).
79. Van Beers, E.H.*, et al.* Intestinal carbamoyl phosphate synthase I in human and rat. Expression during development shows species differences and mosaic expression in duodenum of both species. *J Histochem Cytochem* **46**, 231-240 (1998).
80. Hoogenraad, N.*, et al.* Inhibition of intestinal citrulline synthesis causes severe growth retardation in rats. *Am J Physiol* **249**, G792-799 (1985).
81. Carraway, K.L., Theodoropoulos, G., Kozloski, G.A. & Carothers Carraway, C.A. Muc4/MUC4 functions and regulation in cancer. *Future Oncol* **5**, 1631-1640 (2009).
82. Huang, C.Y. & Tan, T.H. DUSPs, to MAP kinases and beyond. *Cell Biosci* **2**, 24 (2012).
83. Abraham, S.M. & Clark, A.R. Dual-specificity phosphatase 1: a critical regulator of innate immune responses. *Biochem Soc Trans* **34**, 1018-1023 (2006).
84. http://www.genecards.org/cgi-bin/carddisp.pl?gene=KCNE3&search=Kcne3.
85. Shoubridge, C.A., Steeb, C.B. & Read, L.C. IGFBP mRNA expression in small intestine of rat during postnatal development. *Am J Physiol Gastrointest Liver Physiol* **281**, G1378-1384 (2001).
86. Beattie, J., Allan, G.J., Lochrie, J.D. & Flint, D.J. Insulin-like growth factor-binding protein-5 (IGFBP-5): a critical member of the IGF axis. *Biochem J* **395**, 1-19 (2006).
87. Hackett, J.A. & Surani, M.A. DNA methylation dynamics during the mammalian life cycle. *Philos Trans R Soc Lond B Biol Sci* **368**, 20110328 (2013).
88. Auclair, G. & Weber, M. Mechanisms of DNA methylation and demethylation in mammals. *Biochimie* **94**, 2202-2211 (2012).
89. Newell-Price, J., Clark, A.J. & King, P. DNA methylation and silencing of gene expression. *Trends Endocrinol Metab* **11**, 142-148 (2000).
90. Bell, J.T. & Spector, T.D. A twin approach to unraveling epigenetics. *Trends Genet* **27**, 116-125 (2011).
91. Lennartsson, A. & Ekwall, K. Histone modification patterns and epigenetic codes. *Biochim Biophys Acta* **1790**, 863-868 (2009).
92. Vasanthi, D. & Mishra, R.K. Epigenetic regulation of genes during development: a conserved theme from flies to mammals. *J Genet Genomics* **35**, 413-429 (2008).
93. Fisher, C.L. & Fisher, A.G. Chromatin states in pluripotent, differentiated, and reprogrammed cells. *Curr Opin Genet Dev* **21**, 140-146 (2011).

Annex 1: Solutions

PBS

For 1 L of PBS 10X, the following salts were weighted

80 g	NaCl
2g	KCl
14.4g	Na_2HPO_4
2.4g	KH_2PO_4

Water was added until a final volume of 1 L and the solution was autoclaved

Tris-HCl

To prepare 1 L of Tris-HCl 1 M, pH 7.5

121.14 g of Tris were weighted and water was added to dissolve the salt. pH was adjusted with HCl and water added to the final volume.

Hybridization buffer

To prepare 50 ml of Hybridization Buffer the following solutions were mixed

12.5 ml	SSC 20X
25 ml	Formamide 100
5 ml	Dextrane Sulfate 50%
1 ml	Denhart 50X
500 µl	tRNA 10 mg/ml
100 µl	Heparin 50 mg/ml
500 µl	EDTA 0.5M

pH was adjusted to 4.5 – 5.0 with Citric Acid

SSC, pH 4.5 - 5.0

To prepare 1 L of SSC 20X, the following salts were added

175.3 g	NaCl
88.2 g	NaCitrate

Water was added until half the volume, pH was adjusted with Citric Acid and water was added for the final volume of 1 L.

Blocking solution

Add Boehring Blocking Reagent 10% for a final concentration of 1% and goat serum for a final concentration of 10%.

Finalize the volume with TBSX

TBSX

To prepare 1 L of TBSX, the following solutions were mixed

2 ml	KCl 1M
30 ml	NaCl 1M
100 ml	Tris-HCl 1 M pH 7.5
1 ml	Triton X-100 10 %

Water was added for the final volume of 1L.

NTMT

To prepare 500 ml of NTMT, the following solutions were mixed:

25 ml	$MgCl_2$
10 ml	NaCl 1M

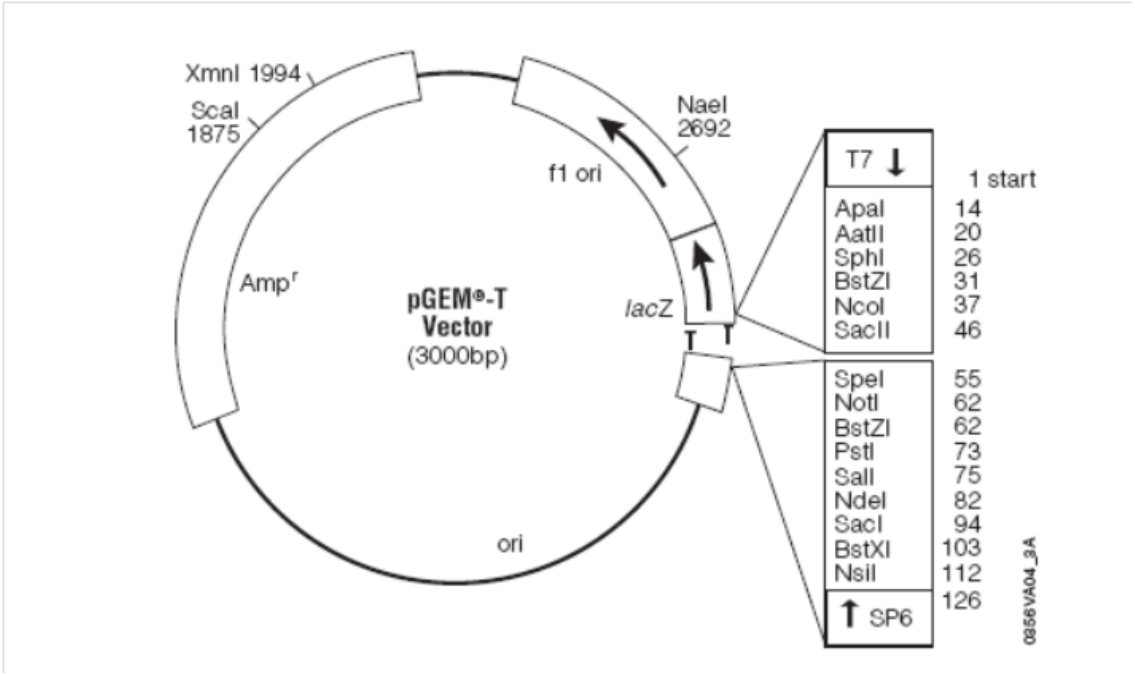
100 ml	Tris-HCl 1 M pH 9.0
500 µl	Tween-20

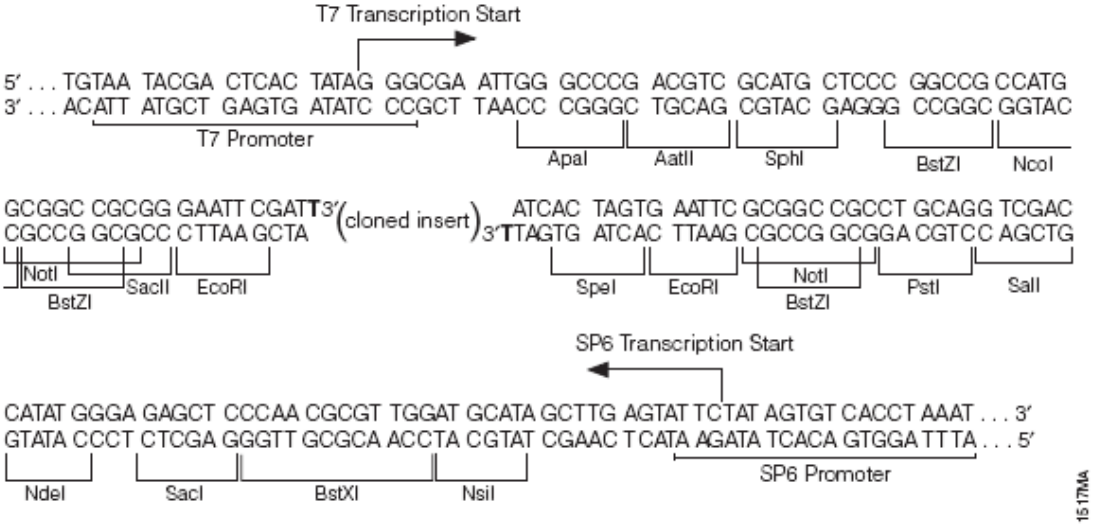
Water was added until the final volume

Staining solution

To the final volume of NTMT (for a final volume of 700 µl per slide), 1 µl/ml of NBT and 3.5 1 µl/ml of BCIP were added.

Annex 2 : pGEM cloning vector map and reference points





Annex 3 : Preparation of competent cells

To obtain competent bacteria cells, one colony of the strain *Escherichia coli* XL-1 Blue, was isolated from a plate of LB medium (30 mg/ml tetracycline) and used to inoculate 5 ml of LB medium. Cells were incubated for 12-16 h at 30ºC and 250 rpm.

1 ml of the culture was inoculated in 250 ml of **SOB** medium and incubated with agitation at 18ºC until a DO of 0.6 nm was reached. The bacterial suspension was incubated on ice to stop growth for 10 min and transferred to 50 ml Falcon tubes and centrifuged 1200 g, 10 min, 4ºC. The supernatant was discarded and the pellet ressuspended in 16 ml of cold transformation buffer (TB), a rich solution of $CaCl_2$, incubated on ice for additional 10 min and centrifuged as described above. DMSO (cryo-preserving agent) was added until a final concentration of 7% was reached and bacteria were incubated for 10 min and aliquoted in 100 µl samples and immediately frozen in liquid nitrogen and stored at -80ºC until further use.

SOB médium (250 mL):

2% Bacto tryptone (5g);

1.5% yeast extract (1.25g);

10 mM NaCl (500 □L stock 5M);

2,5 mM KCl (625 □L stock 1M);

10 mM $MgCl_2$ (2.5 mL stock 1M);

10 mM $MgSO_4$ (2.5 mL stock 1M);

Autoclave after prepare.

TB buffer (100 mL):

10 mM Pipes (0.30 g);

15 mM CaCl2 (1.5 mL stock 1M);

250 mM KCl (25 mL stock 1M);

Adjust pH to 6.7 with KOH 1M, other wise the pipes will not dissolve (take in consideration that the pH changes very fast when near the optimum). After is fully dissolved and the pH is at 6.7 add 55mM $MnCl_2$ (0.69 g). Filter with 0.22 □m filter to sterilize. Do not autoclave the solution. Store it at 4 ºC until use.