

Aus der Klinik und Poliklinik für Hals-, Nasen- und Ohrenheilkunde  
der Ludwig-Maximilians-Universität München

Direktor: Prof. *Dr. med.* Martin Canis

Leiter der HNO Forschung: Prof. *Dr. rer. nat.* Olivier Gires

---

# “Spatiotemporal patterning of EpCAM in murine embryonic differentiation”

---

Dissertation  
zum Erwerb des Doktorgrades der Medizin  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München



vorgelegt von

**Yuanchi Huang**

aus China

2020

Mit Genehmigung der Medizinischen Fakultät der Universität München

Berichterstatter: Prof. Dr. *rer. nat.* Olivier Gires

Mitberichterstatter: Prof. Dr. *rer. biol. hum.* Dorit Nögler

Prof. Dr. *med.* Martin Angele

Prof. Dr. *Dr. h.c.* Christian Haass

Dekan: Prof. Dr. *med. dent.* Dr. Reinhard HICKEL

Tag der mündlichen Prüfung: 09.07.2020

# Eidesstattliche Versicherung

Huang, Yuanchi

---

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Titel

**“Spatiotemporal patterning of EpCAM in murine embryonic differentiation”**

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Munich, 12.07.2020

---

Ort, Datum

Huang, Yuanchi

---

Unterschrift

# CONTENTS

1. INTRODUCTION .....	6
1.1 Early embryonic development of the mouse .....	7
1.1.1 From fertilization to the implantation of a blastocyst.....	7
1.1.2 Gastrulation.....	8
1.2 Stem cells .....	9
1.2.1 Embryonic stem cells.....	9
1.2.2 Adult stem cells .....	10
1.2.3 Stem cell markers .....	11
1.3 Epithelial Cell Adhesion Molecule.....	12
1.3.1 The <i>EPCAM</i> gene .....	13
1.3.2 The structure of the EpCAM protein.....	14
1.3.3 Regulated intramembrane proteolysis (RIP) of EpCAM protein .....	14
1.3.4 Molecular functions of EpCAM.....	16
1.3.5 EpCAM related diseases .....	17
1.3.6 EpCAM in stem cells .....	18
1.4 Objective.....	18
2. MATERIAL.....	20
2.1 Chemicals, consumables, and equipment.....	20
2.1.1 Chemicals and kits.....	20
2.1.1.1 Used chemicals .....	20
2.1.1.2 Used kits.....	21
2.1.2 Consumables .....	21
2.1.3 Equipment .....	22
2.2 Media and Buffers .....	22
2.2.1 Cell culture buffers and media .....	22
2.2.2 Flow cytometry .....	22
2.2.3 Membrane assay.....	23
2.2.4 SDS-PAGE and Western Blot (WB).....	23
2.3 Antibodies .....	23
2.3.1 Primary antibodies .....	23
2.3.2 Secondary antibodies.....	24
2.4 Oligonucleotides .....	25
2.4.1 qRT-PCR primer.....	25
2.4.2 Plasmids .....	26
2.5 Software.....	26
2.6 Cell line .....	26
3. METHODS.....	28
3.1 Cell culture.....	28
3.1.1 Cell culture conditions .....	28
3.1.2 Freezing and thawing of cells.....	28
3.1.3 Cultivation of cells .....	28
3.1.4 Cell counting .....	29
3.1.5 <i>In vitro</i> differentiation of ESCs in embryoid bodies (EBs).....	29
3.1.6 Transfection of ESCs by nucleofection.....	30
3.1.7 Generation of stable cell lines .....	30
3.2 Flow cytometry.....	30
3.3 Biochemical methods.....	31
3.3.1 Preparation of whole cell lysates .....	31
3.3.2 Determination of protein concentration.....	31
3.3.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) .....	32
3.3.4 Immunoblotting (Western blot).....	32
3.3.5 Immunoprecipitation .....	32
3.4 Molecular methods.....	33

3.4.1 RNA concentration measurement .....	33
3.4.2 Isolation of RNA and synthesis of cDNA .....	33
3.4.3 Real Time Quantitative Polymerase Chain Reaction (qPCR) .....	34
3.5 Cell labeling and staining methods .....	35
3.5.1 Immunofluorescence .....	35
3.5.2 Immunohistochemistry .....	36
3.5.3 Laser Scanning Confocal Microscopy .....	36
3.5.4 Epifluorescence microscopy .....	37
3.6 Statistical analysis .....	37
4. RESULTS .....	38
4.1. EpCAM expression and function in mouse ESCs .....	38
4.1.1 Generation of embryoid bodies <i>in vitro</i> .....	38
4.1.2 Characterization of early 3D-differentiation of ESCs .....	39
4.1.3 EpCAM patterning in early 3D-differentiation of ESCs .....	42
4.1.4 EpCAM expression in ecto-, meso-, and endoderm .....	44
4.1.4.1 EpCAM expression in endoderm .....	44
4.1.4.2 EpCAM expression and mesoderm .....	46
4.1.4.3 EpCAM and ectoderm .....	47
4.1.5 Function of EpCAM in ESCs differentiation .....	48
4.1.5.1 Characterization of CRISPR/Cas9-mediated EpCAM knockout ESCs clones .....	49
4.1.5.2 Differentiation of ET14TG2 $\alpha$ WT cells and CRISPR/Cas9 derivatives .....	54
4.1.5.3 Impact of EB size on spontaneous differentiation of ESCs .....	55
4.1.5.4 Mesp1 is important for cardiomyocyte differentiation .....	58
4.1.5.5 EpCAM knockout impacts on cardiomyocyte differentiation .....	61
4.1.6 EpCAM regulates ESCs differentiation via the ERas/AKT cascade .....	63
4.1.6.1 Characterization of ERas as an EpCAM-binding partner .....	63
4.1.6.2 Function of EpCAM/ERas/Akt signaling in ESCs .....	65
4.2 Timing of EpCAM regulation by RIP .....	68
4.2.1 Establishment of an EpCTF-YFP cell model to study the pace of intramembrane proteolysis of EpCAM .....	68
4.2.2 Biochemical evaluation of the cleavage of EpCTF via $\gamma$ -secretase .....	72
4.2.3 The pace of proteolysis of EpCTF is dictated by $\gamma$ -secretase .....	78
4.2.4 Molecular basis for the observed similarity in cleavage pace of EpCTFs from different species .....	82
4.2.5 Assessment of endogenous EpCTF cleavage pace .....	83
4.2.6 EpICD degradation by the proteasome shows high efficiency .....	85
5. DISCUSSION .....	87
5.1 EpCAM is required for full differentiation of pluripotent ESCs .....	87
5.2 Mandatory segregation of EpCAM <sup>+</sup> and EpCAM <sup>-</sup> clusters during EBs differentiation .....	88
5.3 Loss of EpCAM expression in pluripotent ESCs inhibits cardiomyocytes formation .....	89
5.4 EpCAM knockout ESCs clones are blocked in differentiation at the Mesp1 <sup>+</sup> stage .....	91
5.5 Participation of the EpCAM/ERas/AKT axis in endo/ mesodermal differentiation .....	91
5.6 Dynamic EpCAM expression in ESCs differentiation and EMT process .....	93
5.7 RIP as a potential regulatory mechanism of the cell surface expression of EpCAM .....	94
5.8 Proteolysis of EpCTF by $\gamma$ -secretase is a slow process .....	97
5.9 Sustained activation of EpCAM signaling <i>via</i> RIP and efficient shutdown by the proteasome .....	98
6. SUMMARY .....	100
7. ZUSAMMENFASSUNG .....	102
8. APPENDIX .....	105
9. REFERENCE .....	107
10. PUBLICATIONS .....	121
11. ACKNOWLEDGEMENTS .....	123

# 1. INTRODUCTION

One of the most intriguing and fascinating field of research in past decades is related to the embryonic development of mammalian species. The masterplan for the generation of highly complex organisms is laid out at the earliest time points of embryonic development. Despite its obvious importance, we currently have a limited knowledge of cellular dynamics and morphogenetic mechanisms of the human embryonic development (Nowotschin and A. K. Hadjantonakis, 2010). This is partly due to the fact that research of human embryo has limitations, such as ethical restrictions and the “14-day rule”, which does not allow the *in vitro* culture of human embryos beyond day 14 *post* fertilization (De Wert and Mummery, 2003; Aach *et al.*, 2017; Pera, 2017).

Therefore, the mouse model is a popular and powerful tool for the research of embryonic development, without the limitations mentioned above (Taft, 2008; Sozen *et al.*, 2018). The study of mouse embryo has already provided important information about general mechanisms of differentiation, as well as mechanisms that are associated with the pathogenesis of diseases, and which can partially explain defects in morphogenesis and organogenesis (Bedell *et al.*, 1997; Bedell, Jenkins and Copeland, 1997). The embryonic development of the mouse is initiated with the fertilization and is strictly correlated with the age of the embryo. Not only the number of precursors from germ layers and the polarity of cells are important, but also the balance between proliferation and differentiation is essential to ensure the correct timeframe of embryogenesis (Takaoka and Hamada, 2012; Blanpain and Simons, 2013).

Unfortunately, the definition of the precise timing of mouse embryogenesis *in vivo* is limited by the fact that fertilization occurs within the oviduct, and is thus not accessible for experimental assessment (Kojima, Tam and Tam, 2014). But in recent years, the progress in genetic manipulation, an enhancement of *ex utero* cultivation, as well as the option to permanently cultivate pluripotent embryonic stem cells (ESCs) *in vitro* have provided new

opportunities and methods to fill these knowledge gaps (Nowotschin and A. K. Hadjantonakis, 2010).

## **1.1 Early embryonic development of the mouse**

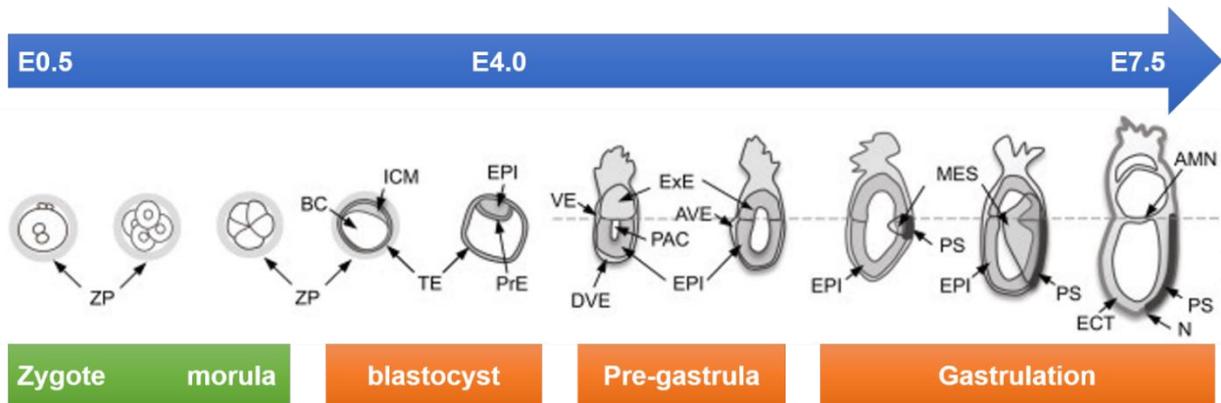
### **1.1.1 From fertilization to the implantation of a blastocyst**

The embryonic age is referred to as days *post coitum* or “Embryonic Day” (embryonic age). As the total cell number of an embryo will be doubled after each cleavage cycle, the embryonic age can be assigned by both cell number and cleavage cycles. As such, the embryo including 2 to 4 blastomeres are at the second cleavage division the embryo with 8 to 16 blastomeres are at the fourth division, and so on (Kojima, Tam and Tam, 2014).

The early development of the mouse embryo starts with fertilization, *i.e.* the merging of the male and female germ cells (Johnson, 2009; Saiz and Plusa, 2013). In this step, the one-cell embryo undertakes cleavage cycles to generate a blastocyst that contains three cell types: the trophectoderm (TE), the epiblast (EPI) and the primitive endoderm (PrE) (**Figure 1**). Subsequently in the pre-implantation stage, cells will be located and segregated to these three lineages at different cleavage divisions (**Figure 1**).

The implantation of a blastocyst into the mother’s uterus represents the next crucial step (BAI *et al.*, 2013) and requires a large number of genetic and cellular interactions (Wang and Dey, 2006). Implantation is composed of three sequential stages: (1) the blastocyst contacts the endometrium (apposition), (2) trophoblast cells attach to the receptive endometrial epithelium (attachment), and (3) trophoblast cells cross the endometrial epithelial (penetration) (Smith, 1980). In the post-implantation period, the inner cell mass (ICM) grows into the blastocyst cavity to form the peri-implantation epiblast (**Figure 1**) (Tam, Williams and Chan, 1993). At this stage the embryo is called egg cylinder and forms a cup-shaped aggregate containing two germ layers, the inner epiblast (EPI) and the outer

visceral endoderm (VE) (**Figure 1**).



**Fig. 1: Scheme of early embryonic development**

Scheme shows features of the mouse embryo at different embryonic ages, from zygote to gastrulation (E0.5–E7.5). Abbreviations: AMN: amnion, AVE: anterior visceral endoderm, BC: blastocyst cavity, DVE: distal visceral endoderm, ECT: ectoderm, EPI: epiblast, ExE: extraembryonic ectoderm, ICM: inner cell mass, MES: mesoderm, N: node, NF: neural fold, PAC: proamniotic cavity, PrE: primitive endoderm, PS: primitive streak, TE: trophectoderm, VE: visceral endoderm, ZP: zona pellucida. Adapted from (Kojima, Tam and Tam, 2014).

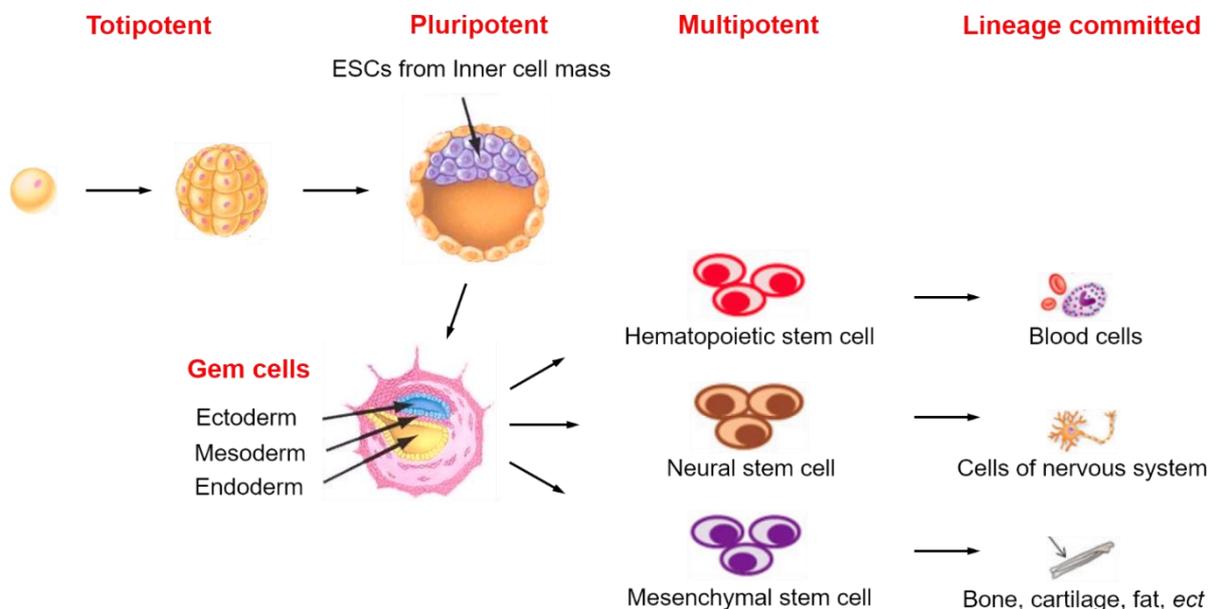
### 1.1.2 Gastrulation

Mouse gastrulation, a constantly evolving 3-dimensional process, transforms two germ layers (EPI and VE) into three layers which are the ectoderm, mesoderm, and endoderm. To achieve this transformation, the primitive streak (PS), a transient embryonic structure, is formed during gastrulation. At this stage, epiblast cells ingress into the junction between the epiblast (ectoderm) and endoderm to form either the mesoderm or the definitive endoderm germ layers. (Parameswaran and Tam, 1995). All three germ layers are partitioned into domains of progenitors, which give rise to specific organs or body parts. During this ingress process, epiblast cells with an epithelial phenotype lose specific cell surface proteins through an epithelial-mesenchymal transition (EMT) and differentiate into mesenchyme (Rivera-Pérez, Mager and Magnuson, 2003; Mikawa *et al.*, 2004; Migeotte *et al.*, 2010; Nowotschin and A.-K. Hadjantonakis, 2010).

## 1.2 Stem cells

Stem cells are defined as toti-, pluri- or multipotent cells which can differentiate into multiple cell types, and have self-renewal capacity to sustain the pool of stem cells. Two major types of stem cells are found in mammalian: ESCs isolated from the ICM of blastocysts, and adult stem cells (ASCs) from various tissues. Although this definition of stemness generally applies to stem cells, it is still necessary to individually characterize embryonic and adult stem cells, because of substantial differences between these cell types.

### 1.2.1 Embryonic stem cells



**Fig.2: Scheme displaying the stem cell hierarchy**

ESCs are pluripotent cells that maintain the ability for differentiating into three germ layers: endoderm, mesoderm, and ectoderm. Adult stem cells (ASCs) such as hematopoietic, neural and mesenchymal stem cells are multipotent stem cells which maintain a more limited capacity for differentiating to specific lineages. Adapted from (Hayes *et al.*, 2012).

ESCs are isolated from the ICM of embryos at the blastocyst stage. The characteristics of ESCs relate to their self-renewal ability *in vitro* and their capacity of differentiating into multiple cell types *in vitro* and *in vivo* (**Figure 2**). Following derivation of the first mouse ESCs in 1981, researchers concentrated on the study of stem cells to define the pluripotency of ESCs (Evans and Kaufman, 1981; Martin, 1981). *In vivo*, mouse ESCs

exhibit a capability of colonizing somatic tissues in diploid aggregation chimeras and support the complete fetal development (Nagy *et al.*, 1990). *In vitro*, the differentiating capacity of ESCs into different tissues is validated by forming embryoid bodies (EBs) that include 3 germ layers (Jaenisch and Young, 2008) (**Figure 2**). Moreover, the ability of ESCs to differentiate into different specific cell types such as cardiomyocytes, liver cells, and osteoblasts is confirmed by several guided differentiation protocols (Buttery *et al.*, 2001; Boheler *et al.*, 2002; Rashidi *et al.*, 2018).

### **1.2.2 Adult stem cells**

Adult stem cells (ASCs) are undifferentiated cells that are found in the body after embryonic development, which bear a potential to differentiate in a lineage-specific manner. As such, ASCs have a more restricted ability of self-renewal and differentiation. Multiple types of ASCs have been isolated from adult tissues, which are responsible for replenishing pools of dead cells (Moore and Lemischka, 2006). Mechanisms to regulate ASCs are strongly associated with their microenvironment, defined as their niche. Niches are locally restricted areas consisting of other cells, extracellular matrix, and signaling factors (Fuchs, Tumber and Guasch, 2004). Moreover, stem cell niches can dynamically regulate the balance between self-renewal and differentiation of ASCs (Morrison and Spradling, 2008).

The best potential for the application of ASCs in clinical trials relies in the restoration (in cell therapy protocols) or replacement (in tissue engineering approaches) of tissues that have been damaged by disease or injury (Pessina and Gribaldo, 2006). Hematopoietic stem cells (HSCs) and mesenchymal stem cells (MSCs) are the most widely used ASCs in this respect. HSCs are used for bone marrow transplantation, while MSCs are attractive for clinical therapy because of their ability to differentiate into specific tissue types, such as bone, cartilage, tendon, and ligament (Beyer Nardi and da Silva Meirelles, 2006). Moreover, adipose-derived processed lipoaspirate cells were identified as a viable clinical alternative to MSCs. (De Ugarte *et al.*, 2003).

Although many studies on the therapeutic potential of ASCs were performed for non-hematologic diseases, there are three major problems: ethical issues, immunological rejection, and the potential of developing teratomas (Moore and Lemischka, 2006).

### 1.2.3 Stem cell markers

in the past decades, different transcriptional factors have been identified as important for the pluripotency status of ESCs. Regulation of these transcriptional factors occurs through a network that can regulate their own expression and that of other key transcriptional factors. A network of three transcriptional factors Sex-determining region Y-Box 2 (Sox2), Octamer 3/4 (Oct3/4) and homeobox transcription factor Nanog plays central functions in the maintenance of pluripotency. These factors continuously inhibit the expression of lineage-specific genes and maintain the expression of pluripotent genes (Boyer *et al.*, 2005; Loh *et al.*, 2006). The working patterns of these factors are not independent to each other, but they are rather involved in an complex network, in which further downstream factors such as *Esrrb*, *Rif1*, and *REST* are implicated as well (Rodda *et al.*, 2005; Loh *et al.*, 2006; Chen *et al.*, 2008; Thomson *et al.*, 2011; Rizzino, 2013).

Apart from transcription factors, various cell surface receptors command the regulation of ESCs pluripotency. For example, three monoclonal antibodies are used to identify stage-specific embryonic antigens (SSEAs) SSEA-1, SSEA-3, and SSEA-4 (Solter and Knowles, 1978; Shamblott *et al.*, 1998). These molecules are involved in cell surface interactions during embryonic development. SSEA-1 is expressed in pre-implantation stage of mouse ESCs, as well as in germ cells. However, it is absent in human ESCs and human embryonic carcinoma cells (Knowles, Aden and Solter, 1978; Solter and Knowles, 1978). Numerous types of surface receptors such as epidermal growth factor (EGF) receptor, leukemia inhibitory factor (LIF) receptor, Epithelial cell adhesion molecule (EpCAM), and bone morphogenetic factor (BMP) receptor are highly expressed in pluripotent ESCs (Bhattacharya *et al.*, 2004; Skottman *et al.*, 2005; Adewumi *et al.*, 2007; Assou *et al.*, 2007;

Lian *et al.*, 2007). Notably, the expression of EpCAM displayed a spatiotemporal regulation in ESCs. EpCAM functions as a growth factor receptor and an adhesion molecule, which shows a high expression in mouse and human ESCs, porcine induced pluripotent stem cells, and in the vast majority of carcinomas (Lu *et al.*, 2010; Ng *et al.*, 2010; Huang *et al.*, 2011; Yu, Ma and Wang, 2017).

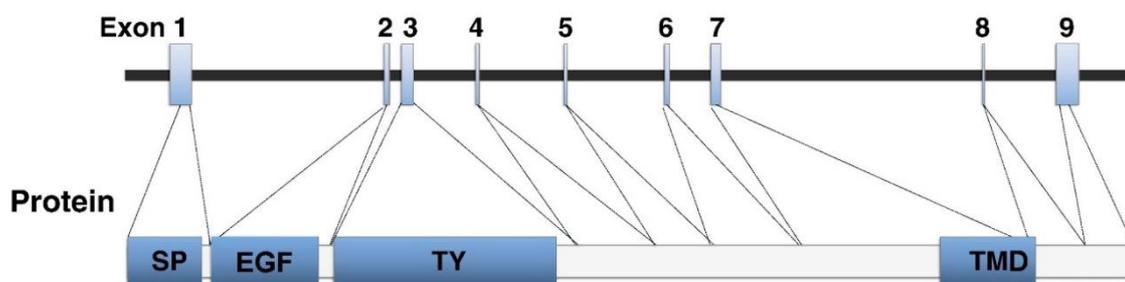
### **1.3 Epithelial Cell Adhesion Molecule**

EpCAM was initially described as a protein on the cell surface, showing a high expression in the majority of carcinomas, including colorectal carcinomas where it was first described, as well as in acute myeloid leukemia (Herlyn *et al.*, 1979; Chadeneau *et al.*, 1991; Bergsagel *et al.*, 1992). Today we reckon the expression pattern of EpCAM as limited to pluripotent ESCs (González *et al.*, 2009; Lu *et al.*, 2010; Ng *et al.*, 2010), hepatic and pancreatic progenitors as well as other endodermal progenitor cells (Schmelzer *et al.*, 2007; Kamimoto *et al.*, 2016; Maimets *et al.*, 2016). Additionally, EpCAM in fully mature cells is expressed in sub-types of normal epithelium composed of adenomatous layers and supra-basal layers of the squamous epithelium (Balzar *et al.*, 1999b), and in malignant cells such as carcinomas and cancer stem cells (Gires, Klein and Baeuerle, 2009; van der Gun *et al.*, 2010). Other matured cell types are completely devoid of EpCAM. This selective expression shows a strong dynamic and tight control of EpCAM expression during the differentiation of ESCs into mature cell types. However, knowledge remains very scarce about the principle and accurate timing of this dynamic expression pattern of EpCAM during ESC differentiation.

### 1.3.1 The *EPCAM* gene

The *EPCAM*-encoding gene, a member of the *GA-733* gene family (Szala et al. 1990, Alberti et al. 1994), is characterized by a high sequence conservation across various species. The human and mouse orthologues of the *EPCAM* gene have a homology of 80 percent (Bergsagel et al., 1992). While human *EPCAM* (*hEPCAM*) localizes on chromosome 2 (chromosomes 2: 47,572,297-47,614,740), mouse *EPCAM* (*mEPCAM*) is encoded on chromosome 17 (Chromosome 17: 87,635,979-87,651,129) (**Figure 3**) (Szala et al., 1990).

**Human chromosome 02: 47,572,297-47,614,740**  
**Mouse chromosome 17: 87,635,979-87,651,129**



**Fig.3: The human and mouse *EPCAM* gene**

The *EPCAM* gene consists of 9 exons. The first exon encodes the signal peptide, exons 2-6 encode the extracellular domain (EpEX), exon 7 encode the transmembrane domain (TMD), and exons 8-9 encode the intracellular domain (EpICD). The EGF (Epidermal Growth Factor)-like domain is encoded by exon 2 while the thyroglobulin domain (TY) is encoded by exon 3. Adapted from (Schnell, Cirulli and Giepmans, 2013).

Both, the mouse and the human *EPCAM* gene consist of nine exons in total. Exons 1-6 encode the extracellular domain of the EpCAM (EpEX), including an epidermal growth factor (EGF)-like domain, a thyroglobulin (TY)-like domain, and a cysteine-depleted region, as well as the signal peptide. The signal peptide (SP) is ultimately cleaved off from the protein, but is essential for the transport of EpCAM to the endoplasmic reticulum and, eventually, to the plasma membrane. The transmembrane domain of EpCAM (TMD) is encoded by exon 7, and the exons 8-9 encode the intracellular domain of EpCAM (EpICD)

**(Figure 3).**

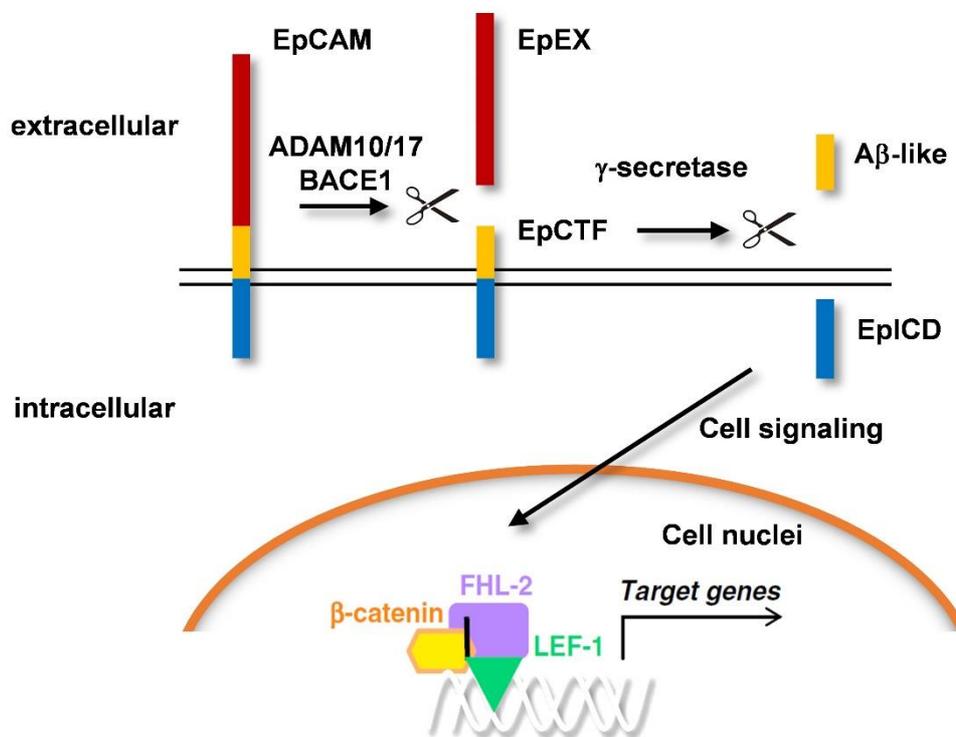
### **1.3.2 The structure of the EpCAM protein**

The mouse EpCAM protein is composed of 315 amino acids (AA), the human EpCAM of 314AA, which are subdivided into three major protein domains: an large extracellular domain of EpCAM (EpEX), a single transmembrane domain (TMD) and a short cytoplasmic domain called EpICD (Balzar *et al.*, 1999a; Munz, Baeuerle and Gires, 2009) **(Figure 4)**. EpEX includes an EGF-like domain and a thyroglobuline (TY) domain. EpEX contains three different N-glycosylation sites at asparagine residues Asn<sup>74</sup>, Asn<sup>111</sup>, and Asn<sup>198</sup> (Thampoe, Ng and Lloyd, 1988; Schön *et al.*, 1993). Glycosylation of Asn<sup>198</sup> contributes to surface expression and protein stability of EpCAM, which was demonstrated by selective mutations of these glycosylation sites (Munz *et al.*, 2008). The crystal structure of the extracellular domain of human EpCAM has been resolved by using standard chromatographic techniques (Pavšič and Lenarčič, 2011). Furthermore, by using mutants of EpCAM within the intracellular domain EpICD, two potential  $\alpha$ -actin binding sites were determined at positions 289 to 296 and 304 to 314, which are essential for EpCAM location at cell-cell adhesion (Balzar *et al.*, 1998).

### **1.3.3 Regulated intramembrane proteolysis (RIP) of EpCAM protein**

Cleavage of EpCAM protein is fulfilled through a complicated process termed regulated intramembrane proteolysis (RIP), which plays an important role in the post-translational processing of numerous transmembrane proteins (Medina and Dotti, 2003; Lal and Caplan, 2011). RIP is involved in initiating cell signaling via surface receptors such as Notch family members (Schroeter, Kisslinger and Kopan, 1998) and EpCAM (Maetzel *et al.*, 2009; Chaves-Pérez *et al.*, 2013).

RIP of EpCAM firstly releases the soluble extracellular domain EpEX by the proteases ADAM10/17 and BACE1. EpEX act as a ligand for intact EpCAM molecules that could induce further cleavage (Maetzel *et al.*, 2009), as well as novel ligand for EGFR in colon and head and neck carcinomas (Liang *et al.*, 2018; Pan *et al.*, 2018). Thereafter, the EpCAM C-terminal fragment (EpCTF) is further processed to an A $\beta$ -like fragments and the intracellular domain EpICD by  $\gamma$ -secretase (Maetzel *et al.*, 2009; Hachmeister *et al.*, 2013; Tsaktanis *et al.*, 2015) (**Figure 4**). Among these fragments, EpICD can form a complex with other molecules and translocate into the cell nucleus (**Figure 4**), which further plays important roles in cancer cell proliferation and stem cell pluripotency maintenance (González *et al.*, 2009; Maetzel *et al.*, 2009; Lu *et al.*, 2010; Huang *et al.*, 2011; Chaves-Pérez *et al.*, 2013). However, pace and efficiency of the generation of EpICD from full-length EpCAM by  $\gamma$ -secretase and the actual stability of EpICD, which is critical for understanding the accurate timing of EpCAM regulation, remain largely unknown.



**Fig. 4: EpCAM undergo regulated intramembrane proteolysis**

EpCAM is firstly cleaved by ADAM10/17, releasing ectodomain (EpEX). Following the first cleavage, EpCAM intracellular part (EpICD) is released by the  $\gamma$ -secretase complex. EpICD form a complex with FHL-2 and  $\beta$ -catenin, then translocate to the cell nuclei. Adapted from (Schnell, Cirulli and Giepmans,

2013).

### 1.3.4 Molecular functions of EpCAM

The name of EpCAM stands for epithelial cell adhesion molecule, which describes a cell-cell adhesion function. Accordingly, EpCAM was initially shown to provide epithelial cells with a weak cell adhesion through homophilic interactions (Litvinov, Bakker, *et al.*, 1994; Litvinov, Velders, *et al.*, 1994). However, the cell adhesion function of EpCAM has been questioned recently and is a matter of debate (Tsaktanis *et al.*, 2015; Gaber *et al.*, 2018). In contrast to E-Cadherin, which involves tight connection via *adherens* junctions, cells that express EpCAM in the absence of classical cadherins are only loosely connected to each other. Notably, when EpCAM is co-expressed in E-Cadherin-expressing cells, it weakened intercellular adhesion mediated by E-cadherin (Litvinov *et al.*, 1997), suggesting complex and intricate functions of EpCAM in cell adhesion.

In vertebrate epithelial cells, the tight junctions (TJs) together with basally localized apical junctions will form the apical junctional complex (AJC). Through electron microscopy, researchers found that *EPCAM*-mutant zebrafish exhibited a reduced expression level of E-cadherin in enveloping layer cells (Huang *et al.*, 2018). Moreover, *EPCAM*-mutant mice displayed a scattered and dispersed TJs in the intestinal epithelia (Lei *et al.*, 2012). All these evidences suggest EpCAM could affect epithelial integrity and lead to dysfunction of AJC and TJs.

RIP of EpCAM induces the release the intracellular domain EpICD and further activates the downstream signaling. EpICD is translocated into the nucleus in a complex that contains additional proteins including four and one-half LIM domains protein 2 (FHL2),  $\beta$ -catenin, and lymphoid enhancer binding factor 1 (Lef-1) (**Figure 4**). In cell nuclei, the EpICD complex can bind to promoters of target genes such as cellular myelocytomatosis oncogene (c-Myc), an oncogenic transcription factor, at Lef-1 consensus sites (Maetzel *et*

*al.*, 2009). In colon carcinoma cell lines, activity of Lef-1 and c-Myc was inhibited after EpCAM knockdown while enhanced expression of EpICD was able to counteract those deficiencies (Münz *et al.*, 2004). Induction of EpCAM expression causes an upregulation of the cell cycle regulating proteins Cyclin A and E, as well as epidermal fatty acid binding protein (Münz, Zeidler and Gires, 2005). Oppositely, repression of EpCAM in squamous cell carcinoma (FaDu) cells lead to a decrease of proliferation (Maetzel *et al.*, 2009; Chaves-Pérez *et al.*, 2013). Similar results were found in breast cancer cells, where inhibition of EpCAM expression led to a reduction in proliferation, migratory and invasive capacity (Osta *et al.*, 2004). Therefore, these evidences suggest an important role of EpCAM signaling in cancer cell proliferation.

### **1.3.5 EpCAM related diseases**

EpCAM plays a role in Lynch syndrome (hereditary nonpolyposis colorectal cancer, HNPCC), an autosomal dominant disorder that predisposes to colorectal adenocarcinoma (60–90%), endometrial carcinoma (20–60%), and various other cancers (Ligtenberg *et al.*, 2009; Kempers *et al.*, 2011; Kuiper *et al.*, 2011). In a total of 45 Lynch syndrome families, 19 deletion mutants of the *EPCAM* gene have been identified (Kuiper *et al.*, 2011). About 20% of Lynch syndrome cases showed a loss of DNA mismatch repair proteins, because mutations in the 3'-end of the *EPCAM* gene result in epigenetic silencing of DNA mismatch repair gene *MSH2* (Rumilla *et al.*, 2011).

Lack of EpCAM was also found in patients suffering from congenital tufting enteropathy (CTE) (Sivagnanam *et al.*, 2008) which is a rare autosomal recessive disease that can cause intractable diarrhea during infancy. CTE patients suffer from chronic diarrhea within the first days after birth and exhibit an impaired growth. At the cellular level, *EPCAM* mutants associated with CTE displayed an abnormal localization within intestinal epithelium and a disorganization of surface enterocytes with focal crowding (Sivagnanam *et al.*, 2008). Several *EPCAM* gene mutations were identified in CTE patients, which can

cause single amino acid exchange, truncation, or partial deletion of the EpCAM protein (Sivagnanam *et al.*, 2008, 2010). A pathological absence of EpCAM may break the required balance between stem cell proliferation and differentiation, which ultimately affects the normal intestinal epithelium development, integrity, and functionality (Sivagnanam *et al.*, 2008).

### **1.3.6 EpCAM in stem cells**

In mouse ESCs, maintenance of self-renewal is associated with a high-level expression of EpCAM. Cultivating mouse ESCs in the absence of LIF, which is required for the maintenance of pluripotency, results in a down-regulation of EpCAM expression together with the expression of c-Myc, Sox2, Oct3/4 and signal transducer and activator of transcription (Stat3) (González *et al.*, 2009). In presence of LIF in the culture medium, repression of EpCAM induced an ESC differentiation, suppressed proliferation, reduced alkaline phosphatase (AP) activity, and Oct3/4, c-Myc, and SSEA-1 expression. In addition, differentiation of mouse ESCs could be partially counteracted by forced expression of EpCAM, which increased Oct3/4 expression and ESCs proliferation, suggesting important functions of EpCAM in ESCs fate regulation (González *et al.*, 2009). In human ESCs, chromatin immunoprecipitation (ChIP) analysis confirmed the direct targeting of the EpICD to promoters of several reprogramming genes, including Oct3/4, Nanog, Sox2, and Krüppel-like factor 4 (Klf4), which could further help maintaining the pluripotency (Lu *et al.*, 2010). Thus, EpCAM is closely associated with the maintenance of the pluripotent state of ESCs.

### **1.4 Objective**

The dynamics and expression patterns of EpCAM has been well studied in cancer cells, and plays important roles in the regulation of cellular fates along the epithelial-mesenchymal transition. However, the regulation of EpCAM expression patterning during

non-pathologic process, *e.g.* during ESCs differentiation and in processes of regeneration of adult cells in organs such as the liver, remains largely elusive. While EpCAM is strongly expressed in pluripotent ESCs, it is only expressed in few mature cell types including primarily epithelia and malignant progeny (*i.e.* carcinomas). Therefore, knowledge of EpCAM regulation and timing of its expression and degradation during ESCs differentiation is required to comprehensively understand the EpCAM expression patterning in non-pathologic processes. Besides classical transcriptional regulation of gene expression and post-translational regulation of protein stability, EpCAM is subject to RIP, which can further feed into its regulation at the protein level. However, the cleavage pace and degradation efficiency of EpCAM, which determine the regulation timing, remain largely unclear.

Therefore, the first goal of the present thesis was to define and characterize the EpCAM expression patterning during ESCs differentiation. The second goal of the thesis was to provide an in-depth characterization of the cleavage pace of EpCTF and of the stability of EpICD.

## 2. MATERIAL

### 2.1 Chemicals, consumables, and equipment

#### 2.1.1 Chemicals and kits

##### 2.1.1.1 Used chemicals

Product	Company
Agarose	Roche, Mannheim
Acrylamide, Protogel ultra-pure	Schröder Diagnostics, Stuttgart
Ammonium persulfate (APS)	BioRad, Hercules
Aqua dest	Braun, Melsungen
$\beta$ -Mercaptoethanol	Sigma-Aldrich GmbH, Taufkirchen
Bovine serum albumin (BSA)	Sigma-Aldrich GmbH, Taufkirchen
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich GmbH, Taufkirchen
DMEM (4.5 g/L glucose)	Biochrom AG, Berlin
EDTA	Carl Roth GmbH & Co.KG, Karlsruhe
Fetal calf serum (FCS)	Biochrom AG, Berlin
Gelatin	Sigma-Aldrich GmbH, Taufkirchen
LightCycler 480 SYBR Green I Master	Roche, Mannheim
Temed	BioRad, Hercules
TissueTek® O.C.T Compound	Sakura Finetek, Staufen
Trypsin/ EDTA	Biochrom AG, Berlin
Tris-(hydroxymethyl)-aminomethan (Tris)	Merck KGaA, Darmstadt
Triton X-100	Sigma-Aldrich GmbH, Taufkirchen
Glycine	Serva GmbH, Heidelberg
Glycerol	Sigma-Aldrich GmbH, Taufkirchen
Paraformaldehyde	Carl Roth GmbH & Co.KG, Karlsruhe
PBS	Pharmacy Klinikum Großhadern, Munich
Propidium iodide	Sigma-Aldrich GmbH, Taufkirchen
Protease Inhibitor Cocktail Complete	Roche, Mannheim
Puromycin	Sigma-Aldrich GmbH, Taufkirchen
ProLong Gold Antifade Mountant	Thermo Fisher Scientific, Waltham
Sodium dodecyl sulfate (SDS)	AppliChem GmbH, Darmstadt
Stempan E14 GMEM	PAN Biotech, Munich

### 2.1.1.2 Used kits

Product	Company
Amaxa™ Mouse ES Cell Nucleofector™ Kit	Lonza Ltd, Basel
BCA Protein Assay	Pierce, Rockford
Immobilon Western Chemiluminescent HRP substrate	Millipore, Bedford
GFP-Trap®	ChromoTek GmbH, Munich
RNeasy Mini Kit	Qiagen, Hilden
QuantiTect Reverse Transcription Kit	Qiagen, Hilden

### 2.1.2 Consumables

Product	Company
Pipettes	Costar, New York
FACS tubes	Becton Dickinson, Heidelberg
Needles Microlance™ 3	Millipore, Schwalbach
Cryotubes	Becton Dickinson, Heidelberg
Pipette tips	American National Can, Menasha
Safe Seal Tips Professional	Starlab, Hamburg
Reagent reservoir	Biozym, Oldendorf
Cell culture flask and dishes	Sarstedt AG & Co., Nümbrecht
Centrifuge tubes 15 mL / 50 mL	Falcon™, New York
Centrifuge tubes 2 mL / 1.5 mL	Sarstedt AG & Co., Nümbrecht
Centrifuge tubes 1.5 mL (nuclease-free)	Eppendorf AG, Hamburg
96-well cell culture plates flat	Costar, New York
96-well ultra-low attachment plates	Corning Incorporated Costar, Amsterdam

### 2.1.3 Equipment

Device	Company
Blotting System Mini trans Blot	BioRad, Hercules
Cell Incubator	Heraeus Holding GmbH, Hanau
Centifuge Mikro 22R	Hettich Lab Technology, Tuttlingen
Centrifuge Rotanta 46 R	Hettich Lab Technology, Tuttlingen
ChemiDoc XRS+ imaging system	BioRad, Hercules
Confocal microscope TCS-SP2	Leica, Bensheim
Flow cytometer FACS Calibur	Becton Dickinson, Heidelberg
Fluorescence microscope Olympus BX43F	Olympus, Tokyo
Light Cycler 480 System	Roche, Mannheim
Magnet stirrer with heat block	Janke & Kunkel, Staufen
Microliter pipettes	Gilson Inc., Middleton
Microplate Reader MRX	Dynatech Laboratories, Bad Nauheim
Microwave	Sharp Electronics GmbH, Hamburg
Multichannel pipette Transferpette-8	Brand GmbH, Wertheim

## 2.2 Media and Buffers

### 2.2.1 Cell culture buffers and media

PBS: 8.0 g NaCl, 0.2 g KCl, 1.15 g Na<sub>2</sub>HPO<sub>4</sub>, 0.2 g KH<sub>2</sub>PO<sub>4</sub> to 1 L H<sub>2</sub>O

Cryopreservation medium: DMEM; 10 % DMSO

DMEM / 10 %FCS: DMEM; 10 % FCS; 1 % Pen Strep

### 2.2.2 Flow cytometry

Flow cytometry (FC) buffer: 3 % FCS in PBS

Antibody solutions: 1:50 in 50 µL FC buffer

Propidium iodide staining solution: 1 µg/ mL propidium iodide (PI) in FC buffer

### **2.2.3 Membrane assay**

Homogenization buffer: 0.2 mL 1M MOPS (pH 7.0), 0.2 mL 1M KCl, 0.2 mL, 100x Complete™ Protease Inhibitor

Assay buffer: 300 µL 0.5 M sodium nitrate, 10 µL 100x complete, 0.5 µL 20 mM ZnCl<sub>2</sub> in 689.5 µL H<sub>2</sub>O

100x complete: 1 complete protease inhibitor tablet in 500 µL H<sub>2</sub>O

Whole cell lysis buffer: 2 complete protease inhibitor tablets, 1% triton-X100 in 10 mL PBS

### **2.2.4 SDS-PAGE and Western Blot (WB)**

Whole cell lysis buffer (2x): 2 complete protease inhibitor tablets, 1% triton-X100 in 50 mL PBS

Laemmli buffer (5x): 62.5 mM Tris pH 6.8, 2% SDS; 10% glycerol, 5% β-mercaptoethanol, 0.001% bromophenol blue

Stacking gel (4%): 13.3 mL 30% acrylamide, 16.6 mL 2 M Tris pH 6.8, 0.663 mL 0.5 M EDTA, 69.44 mL H<sub>2</sub>O

Resolving gel (15%): 50 mL 30% acrylamide, 16.6 mL 2 M Tris pH 8.9, 0.663 mL 0.5 M EDTA, 32.74 mL H<sub>2</sub>O

Running buffer SDS-PAGE: 150 g Tris, 720 g glycine, 50 g SDS to 5L H<sub>2</sub>O

WB washing buffer: 10 tablets PBS, 5 mL Tween-20 to 5L H<sub>2</sub>O

1M Tris Buffer: 121.14 g of Tris base in 800 mL H<sub>2</sub>O

## **2.3 Antibodies**

### **2.3.1 Primary antibodies**

<b>Antibody</b>	<b>Species</b>	<b>Company</b>
FITC anti-Actin IgG <sub>1</sub>	mouse, monoclonal	Santa Cruz, Dallas
anti-EpCAM (Ber-EP4) IgG <sub>1</sub>	mouse, monoclonal	Dako Deutschland GmbH, Hamburg
anti-EpCAM (C-10) IgG <sub>1</sub>	mouse, monoclonal	Santa Cruz, Dallas
anti-EpCAM	mouse, monoclonal	Cell Signaling Technology, Cambridge
(VU1D9) IgG <sub>1</sub> anti-EpICD	guinea pig, polyclonal	Peptide Specialty Laboratories, Heidelberg
anti-GFP/YFP IgG <sub>2a</sub>	mouse, monoclonal	Santa Cruz, Dallas
FICT isotype mouse IgG <sub>1</sub>	mouse, monoclonal	Diatech, Jesi
anti- Flag, mouse IgG <sub>1</sub>	mouse, monoclonal	Sigma-Aldrich GmbH, Taufkirchen
Anti- SSEA1, mouse IgM	mouse, monoclonal	Santa Cruz Biotechnology, Heidelberg
Anti- Vimentin, Rabbit IgG <sub>1</sub>	Rabbit, polyclonal	Abcam, Cambridge

### 2.3.2 Secondary antibodies

<b>Antibody</b>	<b>Company</b>
488-conjugated goat-anti-mouse	Mobitec, Göttingen
IgG Biotinylated horse-anti-mouse IgG (H&L)	Vector Laboratories, Burlingame
Biotinylated goat-anti- rabbit IgG (H&L)	Vector Laboratories, Burlingame
Biotinylated rabbit-anti-rat IgG (H&L)	Jackson ImmunoResearch, West Grove
FITC goat-anti-mouse IgG	Dako Deutschland GmbH, Hamburg
PO rabbit-anti-guinea pig IgG	Sigma-Aldrich GmbH, Taufkirchen
PO goat-anti-mouse IgG	Dako Deutschland GmbH, Hamburg

## 2.4 Oligonucleotides

### 2.4.1 qRT-PCR primer

Primer	Sequence (in 5`-3` orientation)
FW_EpCAM	GCAGCTCAGGAAGAATGTG
BW_EpCAM	CAGCCAGCTTTGAGCAAATGAC
FW_Gusb	CAACCTCTGGTGGCCTTACC
BW_Gusb	CTAGGCATTAAGCTCTCTGTGGATCTCACC
FW_Nkx-2.5	GACAAAGCCGAGACGGATGG
BW_Nkx-2.5	CTGTGCTTGCACCTTGTAGC
FW_Myod1	CCACTCCGGGACATAGACTTG
BW_Myod1	AAAAGCGCAGGTCTGGTGAG
FW_Mesp1	GCTCGGTCCCCGTTTAAGC
BW_Mesp1	ACGATGGGTCCCACGATTCT
FW_Sox2	GACAGCTACGCGCACATGA
BW_Sox2	GGTGCATCGGTTGCATCTG
FW_OCT3/4	CGGAAGAGAAAGCGAACTAGC
BW_OCT3/4	ATTGGCGATGTGAGTGATCTG
FW_Nanog	TCTTCCTGGTCCCCACAGTTT
BW_Nanog	GCAAGAATAGTTCTCGGGATGAA
FW_Eras	TGCCTACAAAGTCTAGCATCTTG
BW_Eras	CTTTTACCAACACCACTTGCAC
FW_aCAA	CTGGATTCTGGCGATGGTGTA
BW_aCAA	CGGACAATTTACGTTTCAGCA
FW_GATA4	CCCTACCCAGCCTACATGG
BW_GATA4	ACATATCGAGATTGGGGTGTCT
FW_aSMA	GGCTGTTTTCCCATCCATCG
BW_aSMA	CAGGCAGTTCGTAGCTCTTCT
FW_wnt5a	CTC TAG CGT CCA CGA ACT CC
BW_wnt5a	CAA ATA GGC AGC CGA GAG AC
FW_wnt11	TTG ACC TGG AGA GAG GTA CAC
BW_wnt11	GTC AGG GGA GCT CTG TAG ATA

## 2.4.2 Plasmids

Plasmid	Description
141pCAG-3SIP	CMV, SV40, IRES, puromycin resistance
141pCAG/YFP	YFP in 141pCAG-3SIP
141pCAG/hEpCAM-YFP	Human EpCAM, YFP tagged in 141pCAG-3SIP
141pCAG/mEpCAM-YFP	Mouse EpCAM, YFP tagged in 141pCAG-3SIP
141pCAG/hEpCTF-YFP	Human EpCTF, YFP tagged in 141pCAG-3SIP
141pCAG/mEpCTF-YFP	Mouse EpCTF, YFP tagged in 141pCAG-3SIP

## 2.5 Software

Software	Company
ApE	Wayne Davis, Salt Lake City
BD Cell Quest Pro Version 5.2.1	Becton Dickinson, Heidelberg
Cell Sense Entry Version 1.8.1	Olympus, Tokyo
Mendeley	Mendeley Ltd., London
GraphPad Prism 7	Graphpad Software Inc., La Jolla
Image Lab Image J	BioRad, Hercules
LAS AF	Leica, Bensheim
LightCycler® 480 SW 1.5	Roche, Mannheim
MS Office 2016	Microsoft, Washington
Adobe reader	Adobe Systems Inc., San Jose
Revelation 4.2.5	DYNEX Technologies Inc., Chantilly

## 2.6 Cell line

The wild type cell line E14TG2 $\alpha$  was isolated from a mouse blastocyst and kindly provided by Markus Conrad (Helmholtz center Munich) and additionally purchased from American Type Culture Collection (ATCC, Manassas, USA). All other cell lines were established by former members of our lab.

Cell line	Description
E14TG2 $\alpha$	Wild type
E14TG2 $\alpha$ - EpCAM-YFP	Transfected with Vector 141pCAG-EpCAM-YFP
E14TG2 $\alpha$ - YFP	Transfected with Vector 141pCAG-YFP
E14TG2 $\alpha$ - Flag ERas	Transfected with Vector 141pCAG-Flag-ERas
E14TG2 $\alpha$ - myr AKT1	Transfected with Vector 141pCAG-myr-AKT1
E14TG2 $\alpha$ - #2	EpCAM+ control for Crispr/cas9
E14TG2 $\alpha$ - #9	EpCAM+ control for Crispr/cas9
E14TG2 $\alpha$ - #16	EpCAM+ control for Crispr/cas9
E14TG2 $\alpha$ - #56	EpCAM Knock out by Crispr/cas9 homozygous
E14TG2 $\alpha$ - #58	EpCAM Knock out by Crispr/cas9 homozygous
E14TG2 $\alpha$ - #62	EpCAM Knock out by Crispr/cas9 homozygous
E14TG2 $\alpha$ - #114	EpCAM Knock out by Crispr/cas9 homozygous
E14TG2 $\alpha$ - #118	EpCAM Knock out by Crispr/cas9 homozygous
E14TG2 $\alpha$ - #138	EpCAM Knock out by Crispr/cas9 homozygous
Bruce4	Wild type
mF9 YFP	Transfected with Vector 141pCAG-YFP
mF9 mEpCAM-YFP	Transfected with Vector 141pCAG-mEpCAM-YFP
mF9 mEpCTF-YFP	Transfected with Vector 141pCAG-mEpCTF-YFP
mF9 hEpCTF-YFP	Transfected with Vector 141pCAG-hEpCTF-YFP
NIH3T3 mEpCTF-YFP	Transfected with Vector 141pCAG-mEpCTF-YFP
HEK293 YFP	Transfected with Vector 141pCAG-YFP
HEK293 hEpCAM-YFP	Transfected with Vector 141pCAG-hEpCAM-YFP
HEK293 hEpCTF-YFP	Transfected with Vector 141pCAG-hEpCTF-YFP
HEK293 mEpCTF-YFP	Transfected with Vector 141pCAG-mEpCTF-YFP
FADU hEpCTF-YFP	Transfected with Vector 141pCAG-hEpCTF-YFP
FADU	Wild type
HCT8	Wild type

# 3. METHODS

## 3.1 Cell culture

### 3.1.1 Cell culture conditions

The passage of mouse embryonic stem cells E14TG2 $\alpha$ , Bruce4, and all other cell culture work was done under a sterile workbench, using sterile tips and solutions.

Cultivation of cells was carried out by default in an incubator at a temperature of 37 °C, a CO<sub>2</sub> content of 5%, and at 95% humidity.

### 3.1.2 Freezing and thawing of cells

Cells were harvested by trypsin treatment and were pelleted at RT for five minutes at 300 g. Subsequently, the cells were dissolved in 1 mL of cold cryopreservation medium, transferred to cryogenic vessels, and slowly frozen at -80 °C before being stored in liquid nitrogen until further use.

After rapid thawing of the cells at 37 °C, DMSO in the cryopreservation medium was diluted by the addition of five volumes of culture medium. The cell pellet obtained by centrifugation (300 g, 5 min.) was then resuspended in fresh medium. After 24 hours, dead cells present in the supernatant were removed by renewing the medium.

### 3.1.3 Cultivation of cells

For the cultivation of mouse ESCs E14TG2 $\alpha$  and Bruce4, cell culture flasks were coated with 0.1% gelatin and subsequently dried 30 minutes under the cell culture workbench. ESCs were cultured in the presence of LIF at a concentration of 0.1 unit /1 mL to maintain the pluripotency of the cells (Smith *et al.*, 1988). mF9 and HEK293 cells were plated in cell culture flasks with DMEM medium containing 20% and 10% FCS respectively.

At 70% confluency, cells were first washed with phosphate-buffer saline solution (PBS)

and then incubated in a 0.5% trypsin solution for five minutes at 37 °C, in order to detach cells from the cell culture surface. After stopping the reaction with FCS-containing medium, cells were centrifuged (300 g, 5 min), resuspended in an appropriate volume of culture medium, and seeded into new cell culture flasks. ESCs were passaged every 48 hours, mF9 and HEK293 cells were passaged every 72 hours.

### **3.1.4 Cell counting**

To determine cell numbers, cells were harvested following trypsin-treatment as described under point 3.1.3. The, 10 µL of the single cell suspension was mixed with trypan blue in a 1:1 ratio and counted using a Neubauer chamber (Thermo Fisher Scientific, Waltham, USA) under a microscope. Trypan blue allows to distinguish dead cells from living cells.

Cells / mL = (cells counted/ number of counted large squares) \* 10<sup>4</sup>.

### **3.1.5 *In vitro* differentiation of ESCs in embryoid bodies (EBs)**

ESCs can be differentiated using the 3D-differentiation hanging drop method (Wang and Yang, 2008). First, ESCs were washed with PBS, harvested by treatment with trypsin, and placed in single cell suspension with differentiation medium (*i.e.* lacking LIF). Subsequently, the total number of cells was determined using trypan blue and a Fuchs-Rosenthal chamber. The cell suspension was diluted to a final concentration of 500 cells for spontaneous differentiation in embryoid bodies (EBs) or 1000 cells for immunohistochemical experiments, each in a volume of 20 µL.

After loading the bottom of a cell culture dish (diameter 20 cm) with 20 mL of PBS, the cell suspension was applied to the lid of the same cell culture dish in drops with a volume of 20 µL using a multichannel pipette. Then, the lid was gently rotated and placed on the dish. After 72 hours at 37 °C, drops containing EBs were transferred to ultra-low attachment 96-well plates containing 160 µL fresh differentiation medium, and were incubated for further 96 hours. Depending on the further procedure, EBs were either maintained in ultra-low

attachment plates or transferred into gelatin-coated, standard 96-well flat bottom plates.

### **3.1.6 Transfection of ESCs by nucleofection**

ESCs were harvested and counted as described in paragraphs 3.1.3 - 3.1.4. For each transfection,  $2 \times 10^6$  cells were centrifuged (300 g, 5 min), resuspended in 100  $\mu$ L Amaxa Nuclei transfection reagent, and added to plasmid DNA in the Nuclei transfection solution (Lonza Cologne AG, Cologne, Germany). ESCs transfection was then carried out in a Nucleofector® (Lonza Cologne AG, Cologne, Germany) using the program A-24 according to the manufacturer's instructions.

### **3.1.7 Generation of stable cell lines**

To generate stable cell lines, ESCs were transfected as described in paragraph 3.1.6. 24h after transfection, puromycin at a concentration of 1  $\mu$ g / 1  $\mu$ L was added to the cell medium to select for resistant cells. Cells were cultivated for several weeks in the presence of puromycin and subsequently analyzed by flow cytometry, western blot, and/or qRT-PCR to ensure the expression of the transfected mRNA and protein in the selected cell population.

## **3.2 Flow cytometry**

To determine the expression of proteins on the cell surface, the method of flow cytometry was applied. For this purpose, primary antibodies that bind the extracellular domain of a protein were used in combination with fluorescein isothiocyanate (FITC)-conjugated secondary antibodies. All antibodies were used in a 1:50 dilution.

The cells to be examined were washed with PBS, detached from culture plates through treatment with trypsin, and harvested in cell culture medium. Subsequently, the cell pellet was washed again in PBS, were resuspended in FACS buffer, and were treated with the primary antibody for 15 minutes at room temperature (RT). As a control for unspecific

binding of secondary antibodies, primary antibodies were omitted. Before staining with secondary antibody for 15 minutes, cells were centrifuged and supernatants removed. After repeated centrifugation, the pellet was resuspended in 500  $\mu$ L FACS buffer. Furthermore, 0.5 mg / mL of propidium iodide (PI) was added to the cell suspension to distinguish between dead and living cells in the following analysis. The assessment and analysis of samples was conducted on a FACScalibur device using the CellQuest™ software (Becton, Dickinson and Company, Franklin Lakes, USA). PI-positive cells were excluded from the analysis. Expression values represent ratios of mean fluorescence intensities of the antigen of interest divided by negative controls.

### **3.3 Biochemical methods**

#### **3.3.1 Preparation of whole cell lysates**

Cells were collected, washed once with PBS, and centrifuged for 5 min at 280 g and RT. The supernatant was discarded and the cell pellet was resuspended in a two-fold volume of 2x whole cell lysis buffer. Alternatively, pellets were frozen at -80 °C for several days before lysis. After lysis, samples were centrifuged for 10 min at 16000 g and 4 °C to collect supernatants which will be transferred into a new tube. BCA-assay will be performed to determine the protein concentration. Laemmli buffer was mixed with samples, and then heated at 95 °C for 5 min before separation on SDS-PAGE. Protein samples were stored at -20 °C.

#### **3.3.2 Determination of protein concentration**

Protein concentrations were detected by using the BCA assay kit. 5  $\mu$ L of the protein samples were mixed with 100  $\mu$ L BCA solution and the absorbance at 595 nm wavelength was measured with a spectrophotometer (GeneQuantPro, GE Healthcare, Chicago, USA). All measurements were performed in triplicates.

### **3.3.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)**

SDS-PAGE is to separate proteins by mass. As a result, proteins with a smaller molecular weight migrate faster than those with a higher molecular weight. Per gel, 10 mL resolving gel (15%) were mixed with 50  $\mu$ L APS and 30  $\mu$ L TEMED. After polymerization, 2 mL of the stacking gel was mixed with 30  $\mu$ L APS and 15  $\mu$ L TEMED, poured, and polymerized on top of the separation gel additionally using a comb to generate loading pockets. Subsequently, same amounts of proteins of whole cell lysate samples were loaded on gels. Gel electrophoresis was conducted for 15 min at 15 mA and 2 h at 30 mA in SDS running buffer. Afterwards, gels were used for immunoblotting.

### **3.3.4 Immunoblotting (Western blot)**

A wet blotting system (BioRad, Hercules, USA) was used in our study. With this system, proteins separated in a polyacrylamide gel will be transferred to a polyvinylidene fluoride (PVDF) membrane. To do so, membranes were first incubated in methanol for 1 min and then transferred into blotting buffer. Blotting was conducted for 50 min at 100 V and RT. After blotting, PVDF membranes were firstly incubated in blocking solution for minimally 30 min at RT. Membranes were washed in washing buffer for 5 min and incubated in primary antibody for 1 h at RT or over-night at 4 °C. Subsequently, membranes were washed three times in PBST for 5 min and incubated with the secondary antibody for 30 min at RT. Antigen-antibody reactions were shown via the chemiluminescent HRP substrate (Millipore, Burlington, Massachusetts, USA). Images were acquired by a ChemiDoc XRS+ imaging system (BioRad, Hercules, USA) and analyzed by ImageLab software.

### **3.3.5 Immunoprecipitation**

Lysates from cell culture were prepared as described in 3.3.1. Vortex GFP-Trap® A beads (ChromoTek GmbH, Munich, Germany) and pipette 25  $\mu$ l bead slurry into 500  $\mu$ l ice-cold

dilution buffer. Centrifuge at 2500 g for 2 min at 4 °C. Discard supernatant and repeat wash twice in cold PBS. Add cell lysate to equilibrated GFP-Trap® A beads, gently rotate for 1 hour at 4 °C. Centrifuge at 2500 g for 2 min at 4 °C and discard remaining supernatant. Resuspend beads in 500 µL ice-cold dilution buffer. Centrifuge at 2500 g for 2 min at 4 °C and discard supernatant, repeat this wash twice. Resuspend beads in 100 µL 2x SDS-sample buffer and boil for 10 min at 95°C to dissociate immunocomplexes from GFP-Trap® A beads. Beads can be collected by centrifugation at 2500 g while supernatant containing the target proteins can be further performed with the SDS-PAGE.

### **3.4 Molecular methods**

#### **3.4.1 RNA concentration measurement**

The concentration of RNA was determined with a Nano drop device (Implen GmbH, Munich, Germany). To determine the concentration of nucleic acids in solution, the absorbance as measured at the wavelength of 260 nm is used. The ratios of 260 nm /280 nm and 260 nm /230 nm are calculated for an indication of the purity of the samples. Pure RNA samples have expected 260/280 ratios of more than 2.0.

#### **3.4.2 Isolation of RNA and synthesis of cDNA**

The total RNA isolation of cells was performed with the RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. For this purpose,  $1 \times 10^6$  cells were harvested, washed with PBS, and digested in a corresponding buffer with QiaShredder columns. Complementary DNA (cDNA) reverse transcription was performed with the QuantiTect Reverse Transcription kit (Qiagen, Hilden, Germany) as described below. To avoid degradation of total RNA, all pipetting steps were performed on ice according to the manufacturer's recommendations. By default, 1 µg of total RNA was reverse transcribed using the two-step protocol from Instruction manual of QuantiTect Reverse Transcription kit. After digestion of residual genomic DNA in the first step by using DNase, the treated

RNA was reverse transcribed to cDNA in the second step. This reaction was terminated by inactivation of the reverse transcriptase at 95 °C for 1 min. A small volume of RNA from the first step was used as a negative control in the following quantitative polymerase chain reaction (qPCR) to exclude RNA contamination with genomic DNA.

### 3.4.3 Real Time Quantitative Polymerase Chain Reaction (qPCR)

For the analysis of gene expression by qPCR, the plate cycler LightCycler® 480 and the double-strand-specific master mix LightCycler® 480 SYBR Green I Master (Roche, Basel, Switzerland) were used. The sample was prepared according to the following table:

Volume	Reagent
5 µL	SYBR Green I Master
1 µL	cDNA
1 µL	Primer sense
1 µL	Primer antisense
2 µL	H <sub>2</sub> O
10 µL	End volume

The reaction mixture and the protocol are summarized in the following table:

Temperature	Time	Ramp rate (°C/s)	Cycles	Detection mode	Reaction
95 °C	10 min	4,4	1	--	Preincubation
95 °C	10 sec	4,4	45	--	Amplification
65 °C	10 sec	2,2		--	
72 °C	15 sec	4,4		"Single"	
95 °C	5 sec	4,4		--	
65 °C	1 min	2,2	1	--	Melting curve
97 °C		0,11		continuous	

The cycle threshold (CT) value is a theoretical value indicating the beginning of the exponential phase of a PCR reaction. The normalization of the CT values of all genes was made with the CT value of the housekeeping gene.

Calculation was as follows:

1. Average of Cp-values:  $C_p = \Sigma C_p / 3$

2. Normalization to a housekeeping gene (HG):  $\Delta C_p = C_p(\text{gene}) - C_p(\text{HG})$

3. Calculation of relative gene expression levels:

a) Control group (was set to "1.0"):  $\Delta\Delta C_p(\text{control}) = 2^{-(\Delta C_p(\text{control}) - \Delta C_p(\text{control}))}$

b) Sample group:  $\Delta\Delta C_p(\text{sample}) = 2^{-(\Delta C_p(\text{sample}) - \Delta C_p(\text{control}))}$

### **3.5 Cell labeling and staining methods**

#### **3.5.1 Immunofluorescence**

Embryoid bodies (EBs) generated as described in paragraph 3.1.5 and cell pellets were placed in cryomolds, covered bubble-free with the embedding medium "TissueTek" (Science Services GmbH, Munich, Germany), snap-frozen with liquid nitrogen, and stored at -20 °C. Using a cryo-microtome, 4µm sections were generated, mounted on Super Frost slides (Thermo Fisher Scientific, Waltham, USA), and dried at RT before being stored again at -20 °C until further use.

EBs were washed with PBS for 5 min and fixed with 3.5% paraformaldehyde (PFA) for 10 min in the dark at 4 °C and 5 min in the dark at RT. EBs were then washed with PBS for 5 min, permeabilized by 4 °C methanol, and blocked with 200 µL horse serum for 20 min at RT. Next, EBs were incubated with the first antibody overnight at 4 °C. After three times washing in PBS for 5 min, EBs were incubated with an Alexa 488-linked secondary antibody 1 h in the dark at RT. Finally, EBs were covered with ProLong™ Gold Antifade Mountant with DAPI (Thermo Fisher Scientific, Waltham, USA) to stain nuclei.

Cells were grown on glass slides and proliferate to 50% confluence followed by treatment with DAPT or β-lactone. Cells were washed with PBS for 5 min, fixed in 3% PFA for 10 min in the dark at RT, cells were then washed in PBS for 5 min and covered with ProLong™ Gold Antifade Mountant with DAPI (Thermo Fisher Scientific, Waltham, USA) to stain nuclei (Vector Laboratories, Burlingame, USA).

Both EBs and cells staining were analyzed using a TCS-SP5 scanning system and the

LAS AF software (Leica, Wetzlar, Germany).

### **3.5.2 Immunohistochemistry**

Samples were fixed in acetone for 5 min at RT, followed by fixation with 3.5% PFA for 10 min in the dark at 4 °C and 5 min in the dark at RT. Next, samples were washed twice in PBS for 5 min at RT and incubated with horse serum for 20 min at RT. Incubation with first antibody was performed for 1 h at RT or over-night at 4 °C. After washing samples with PBS and Brij solution (Thermo Fisher Scientific, Waltham, USA), sections were incubated with a biotinylated anti-mouse antibody for 30 min at RT, washed again with PBS and Brij solution, and subsequently incubated with a peroxidase-labeled avidin–biotin complex. Finally, samples were stained with amino-ethylcarbazole (AEC), generating a red-brown staining of the antigen/antibody complexes. Counterstaining was achieved with hematoxylin (blue). Samples were covered with Kaiser's glycerol gelatin and images were acquired via an Olympus BX43F fluorescence microscope and CellEntry software (Olympus, Tokyo, Japan).

### **3.5.3 Laser Scanning Confocal Microscopy**

For confocal microscopy analysis, EBs and cells were prepared as described in 3.5.1. Images were acquired with a TCS-SP5 laser scanning system, 63x oil immersion objective, in which three major filters at wavelengths 358 nm (DAPI), 488 nm (GFP), and 647 nm (red) were used to capture the images. The microscope was initially switched to the laser scanning live mode at DAPI filter to locate the cells, then the images from green and red channel will be recorded at imaging acquiring mode. After these settings were determined, the configuration was saved and used for acquisition and analysis of all images in the study from different samples. All images were analyzed by LAS AF software (Leica Microsystems; Wetzlar, Germany).

### **3.5.4 Epifluorescence microscopy**

Cells were grown to 50% confluence at 10 cm tissue culture dishes. DAPT was added into medium 24 hours before the experiment to block the function of  $\gamma$ -secretase. After 5 times PBS wash, new medium without DAPT was added to cells. Images were acquired with an Olympus motorized inverted research microscope IX81 and MMI Cellcut Plus software (Molecular Machines & Industries, Eching, Germany). Fluorescence intensity across the plasma membrane areas was quantified with Fiji software. 959ms in 200x magnification and 1180ms in 400x magnification were set as exposure time respectively. 0.5 dB was set as gain value and image background was subtracted from all images before quantification. Mean fluorescence values at time- point 0 was set to 100% as a reference, all values from subsequent time points were normalized to the reference.

### **3.6 Statistical analysis**

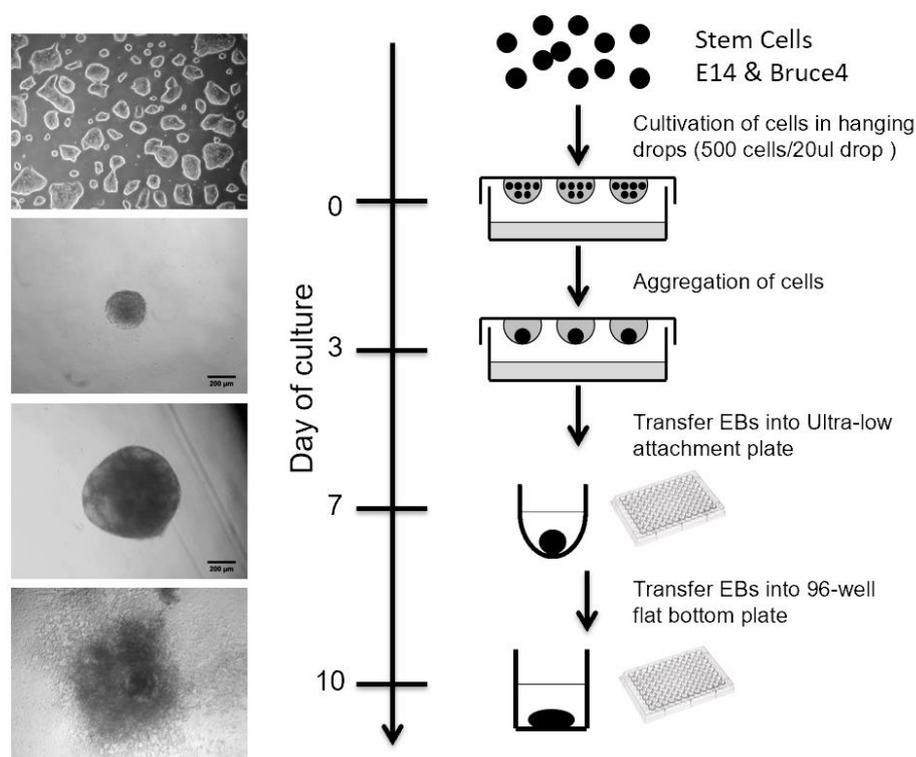
Results are presented as mean value  $\pm$  standard error of the mean (SEM) of  $\geq 3$  independent experiments unless indicated otherwise. Significant differences between two groups were calculated by a Student's T-test in Prism (GraphPad Software, San Diego, USA). Significant differences of more than two groups were calculated with ANOVA test with Bonferroni corrections in Prism. Levels of significance were displayed as \*p-value < 0.05; \*\*p-value < 0.01; \*\*\*p-value < 0.001; \*\*\*\*p-value < 0.0001, and referred to control group unless depicted otherwise.

# 4. RESULTS

## 4.1. EpCAM expression and function in mouse ESCs

### 4.1.1 Generation of embryoid bodies *in vitro*

ESCs are capable of generating the three germ layers ecto-, meso-, and endoderm *in vitro*, and thereby can help understanding early embryogenesis at the cellular and molecular levels (Nishikawa, Jakt and Era, 2007). Here, a hanging-drop 3D-differentiation model was used to generate embryoid bodies (EBs) from E14TG2 $\alpha$  and Bruce4 ESCs (**Figure 5**). EBs closely mimic embryogenesis *in vitro* and allow genetic manipulations of the cells of interest. The method used herein allows to generate EBs from a defined cell number and with reproducible size, in which spontaneous differentiation occurs *in vitro*.



**Fig. 5: Hanging drop method**

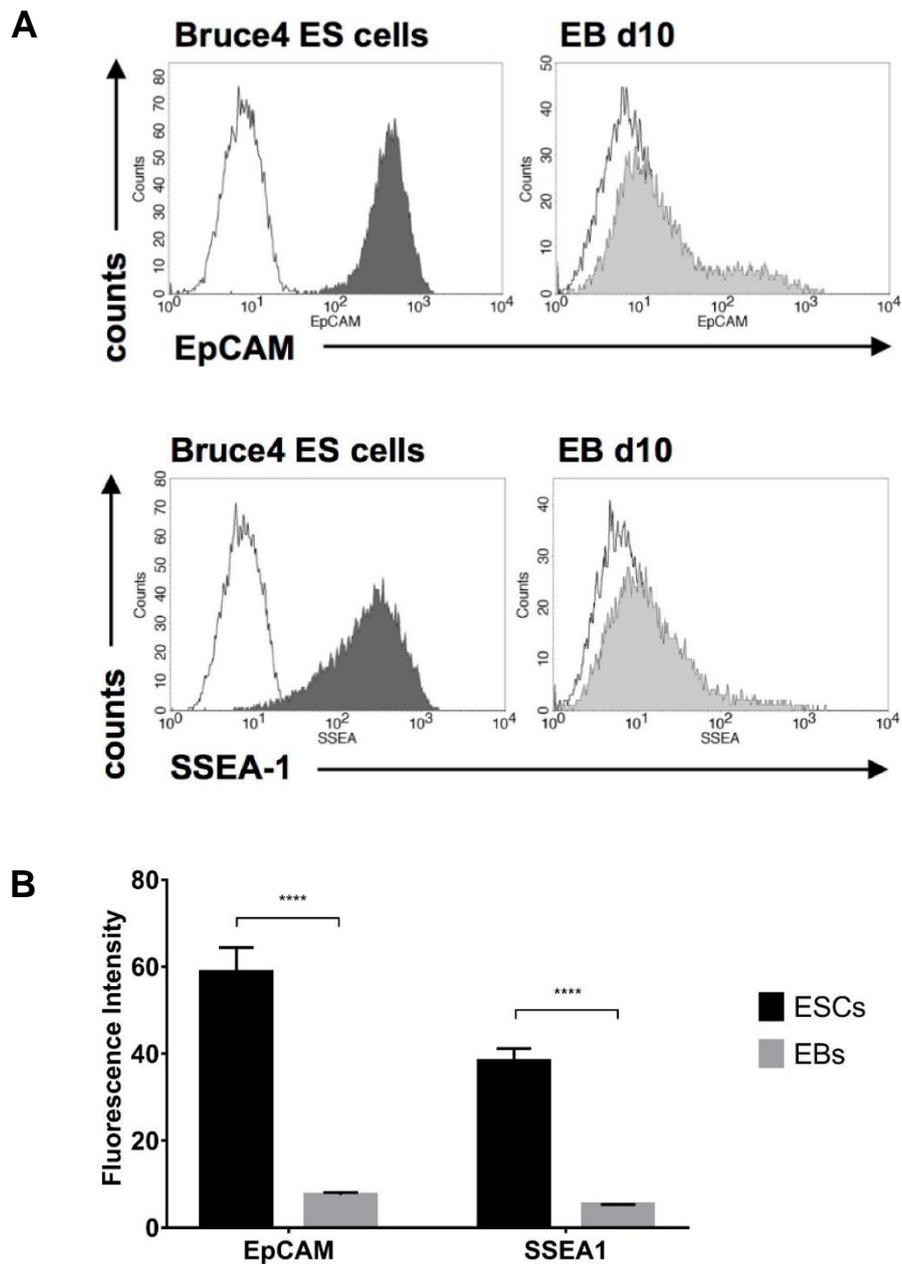
Pluripotent E14TG2 $\alpha$  or Bruce4 ESCs were harvested, placed in single cell suspension, and cultured using the hanging-drop method (500 cells per drop in 20  $\mu$ L volume). After three days of incubation, the spheroid-like structures termed embryoid bodies (EBs) were transferred to ultra-low attachment multi-well plates for an additional four days. On day seven, EBs were transferred to gelatin-coated plates, in

which they were incubated until the end of each experiment. Wild-type cells were photographed on different days with a microscope. While the first picture shows the morphology of cells initially used for the generation of EBs, the following two non-adherent EBs are depicted. On day ten, the EBs were already attached to the bottom surface of the plate and was further differentiated. Shown are representative pictures of the respective culture conditions of the EBs.

**Figure 5** shows the exact procedure of the hanging-drop method. A single cell suspension consisting of  $2.5 \times 10^3$  cells / mL in differentiation medium is prepared and 20  $\mu$ L drops are placed on the lid of a 20 cm diameter cell culture dish for three days in an incubator under standard conditions. On day three, EBs formed in hanging drops are transferred into ultra-low attachment multi-well plates, to which EBs do not adhere. EBs will be then incubated for additional four days at 37 °C and subsequently transferred into gelatin-coated 96-well culture plates to be kept for different durations, depending on the particular experimental setup (**Figure 5**).

#### **4.1.2 Characterization of early 3D-differentiation of ESCs**

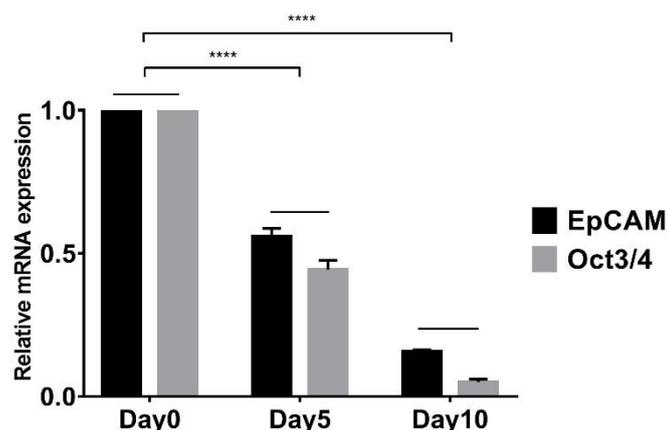
The formation and spontaneous differentiation of EBs from E14TG2 $\alpha$  ESCs were already evaluated in our lab (Sarrach *et al.*, 2018). In the present study, EBs were generated from a second ESC line, namely Bruce4, and were cultured for 10 days in differentiation medium lacking LIF. Differentiation of the cells within EBs was verified by staining of the pluripotency cell surface marker SSEA-1 (Solter and Knowles, 1978; Williams *et al.*, 1988) after disintegration of EBs into single cell suspension (**Figure 6A**). Additionally, EpCAM on the cell surface was also examined in a pluripotent state and at day 10 of differentiation by flow cytometry.



**Fig. 6: EpCAM and SSEA-1 expression on pluripotent and differentiated Bruce4 ESCs**

(A) EBs were generated from Bruce4 ESCs and were cultured for 10 days in differentiation medium lacking LIF. The expression of EpCAM and pluripotency marker SSEA-1 on the surface of ESCs (dark grey) and differentiated EBs (light grey) were determined by flow cytometry using EpCAM- and SSEA-1-specific antibodies. Negative controls are displayed as black lines. Shown are representative histograms from  $n = 3$  independent experiments. (B) Shown are the means and standard deviations of EpCAM and SSEA-1 expression in pluripotent ESCs and EBs from  $n = 3$  independent experiments.  $p$ -value were calculated with a Two-way ANOVA test and multiple *posthoc* comparisons with Bonferroni correction. \*\*\*\* $<0.0001$ .

**Figure 6A** shows representative histograms of the flow cytometry measurement of the membrane proteins EpCAM and SSEA-1. EpCAM is highly expressed in ESCs in the pluripotent state, while it is strongly reduced in EBs after 10 days of differentiation (**Figure 6A**, top). A similar expression pattern of SSEA-1 was assessed by flow cytometry in the same samples (**Figure 6A**, bottom). Mean cell surface expression levels of EpCAM and SSEA-1 were calculated from three independent experiments, and revealed 87% and 86% decreases after 10 days of differentiation, respectively (**Figure 6B**). The reduction of SSEA-1 at the protein level confirmed a differentiated state of Bruce4 ESCs after 10 days of 3D-differentiation in EBs. Furthermore, a subpopulation of differentiated ESCs retained the expression of EpCAM, as visualized by a “shoulder” of EpCAM<sup>+</sup> cells in flow cytometry histograms (**Figure 6A**).



**Fig. 7: EpCAM and Oct3/4 expression during differentiation of Bruce4 ESCs**

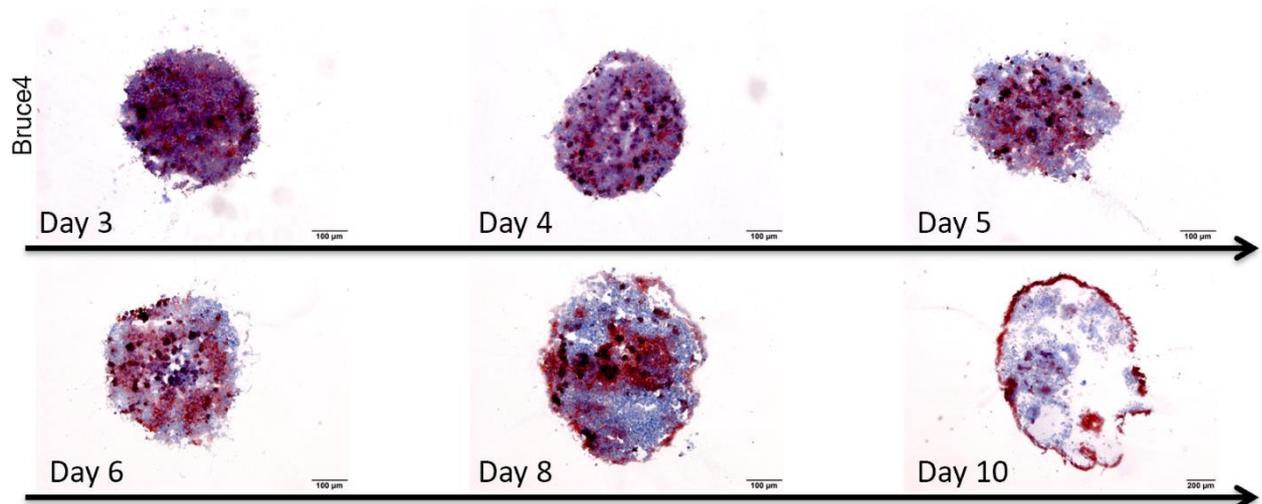
Bruce4 ESCs were differentiated in EBs for 10 days. Bruce4 ESCs were then harvested on day 0, 5, and 10 of differentiation, RNA was isolated, cDNA was synthesized, and the expression of EpCAM and Oct3/4 was quantified by quantitative RT-PCR. The values measured in n = 3 independent experiments were averaged and presented with the corresponding standard deviations. p-value were calculated with a Two-way ANOVA test and multiple *posthoc* comparisons with Bonferroni correction. \*\*\*\*<0.0001.

Next, Bruce4 ESCs were cultured for 10 days and harvested on day 0, 5, and 10 of differentiation. RNA was isolated at each time point and quantitative RT-PCR (qRT-PCR) was performed to evaluate the expression of Oct3/4 at these three time points. The results revealed a substantial reduction of the pluripotency marker Oct3/4 at the transcriptional

level over the time (**Figure 7**, grey bar). On day five, expression of Oct3/4 was decreased to 44%, while on day 10, it was decreased to 5% of the initial levels of undifferentiated Bruce4 ESCs. Similar to the reduction of the pluripotency marker Oct3/4, transcription of the *EPCAM* gene was reduced to 45% on day five and to 15% on day 10, as compared to the levels of *EPCAM* in undifferentiated, pluripotent Bruce4 ESCs (**Figure 7**, black bar). Hence, Bruce4 ESCs could be reproducibly differentiated in EBs, with a loss of SSEA-1, Oct3/4, and EpCAM expression that was comparable to levels and kinetics observed in E14TG2 $\alpha$  ESCs (Sarrach *et al.*, 2018).

#### 4.1.3 EpCAM patterning in early 3D-differentiation of ESCs

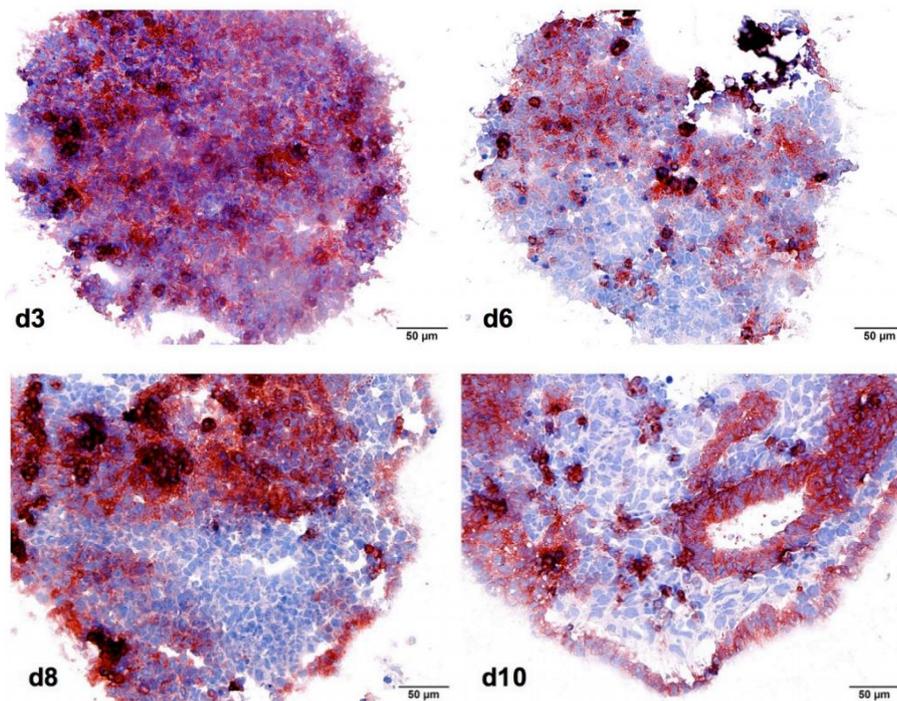
Heterogeneous down-regulation of EpCAM was observed in previous flow cytometry results, with a majority of cells losing the expression and a minor population retaining EpCAM. Therefore, in order to further understand the regulation of EpCAM patterning during early differentiation of EBs, pluripotent Bruce4 ESCs were plated in hanging drops and EBs were harvested at different time points, cryo-sections were generated, and were then stained with an EpCAM-specific antibody.



**Fig. 8: EpCAM patterning in early 3D-differentiation of Bruce4 EBs**

Bruce4 ESCs were harvested, washed, and plated in hanging drops and transferred after three days to ultra-low attachment plates. From day three on, EBs were shock-frozen, sectioned, and stained with EpCAM-specific antibody for the time points of day 3, 4, 5, 6, 8 and 10. Shown are representative staining at each time point from  $n = 3$  independent experiments with 18 EBs in each experiment.

In differentiating EBs from Bruce4 cells, loss of EpCAM expression and segregation of clusters of EpCAM<sup>+</sup> and EpCAM<sup>-</sup> cells occurred from day four to five onwards, and resulted in spatiotemporal patterning of EpCAM, eventually resulting in an outer margin area composed of EpCAM<sup>+</sup> cells, while the majority of cells within the EBs had lost EpCAM (Figure 8).



**Fig. 9: Selective expression of EpCAM in differentiated ESCs in EBs**

Bruce4 ESCs were harvested, washed, and plated in hanging drops, and transferred after three days to ultra-low attachment plates. Subsequently, the EBs from different time points at day 3, 6, 8, and 10 were shock-frozen, sectioned, and were stained with EpCAM-specific antibody. Shown are the representative staining at each time points from n = 3 independent experiments with 12 EBs in each experiment.

Further enlargements in **Figure 9** show that EpCAM<sup>+</sup> cells were homogeneously distributed in EBs at day three of differentiation. At day six of differentiation, a distinct patterning of EpCAM expression became obvious. From day eight onwards, a minor population of cells with strong EpCAM expression was localized at the margin of the EBs and linings of vacuoles. At day 10 of differentiation, the majority of cells revealed EpCAM<sup>-</sup> (Figure 8, 9). Hence, EpCAM displays a spatiotemporally selective expression throughout

spontaneous ESCs differentiation.

#### **4.1.4 EpCAM expression in ecto-, meso-, and endoderm**

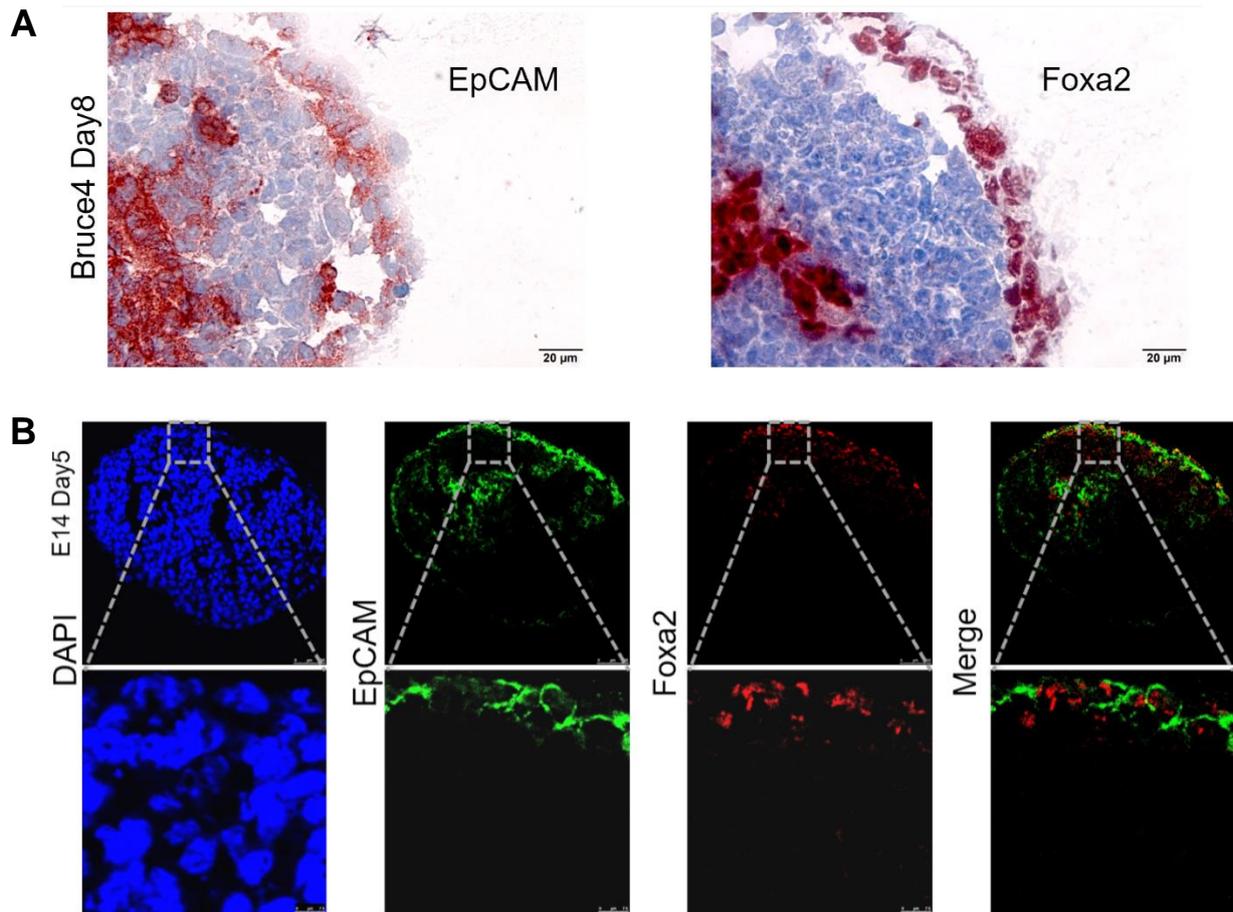
During 3D-differentiation of Bruce4 ESCs, a selective loss and maintenance of EpCAM expression was observed in EBs. EBs are comprised of three embryonic germ layers, *i.e.* ecto-, meso-, and endoderm (Itskovitz-Eldor *et al.*, 2000). To investigate the correlation between the selective expression of EpCAM and the three germ layers in EBs, immunohistochemistry (IHC) and immunofluorescence (IF) staining were performed on section of EBs by using antibodies specific for each germ layer.

##### **4.1.4.1 EpCAM expression in endoderm**

Foxa2 is a member of a family of nuclear transcription factors that play a role in cell commitment, differentiation, and organ-specific gene transcription. Foxa2 also has an important function in the regulation of epithelialization in mouse endoderm (Burtscher and Lickert, 2009a). Here, we used Foxa2 as an endoderm-specific marker to study the distribution of EpCAM within endodermal cells in differentiated ESCs in EBs.

Consecutive sections of Bruce4 EBs were stained with EpCAM and Foxa2, respectively, at day eight of EBs differentiation. At this time point, patterning of EpCAM is pronounced. EpCAM expression partially overlapped with Foxa2 expression. A similar localization of EpCAM with Foxa2 was primarily observed in marginal cells in consecutive sections of EBs, which could represent cells of the visceral endoderm, and in more central areas of EBs (**Figure 10A**). To further analyze the correlation of EpCAM and Foxa2, immunofluorescence double-staining of EpCAM and Foxa2 was performed with E14TG2 $\alpha$  EBs. In differentiated E14TG2 $\alpha$  EBs at day 5, EpCAM and Foxa2 co-localized in cells at the edge of EBs. Magnification of fluorescence staining confirmed at the cellular level that EpCAM was localized on the cell surface, while Foxa2 was expectedly expressed

within cells (**Figure 10B**).

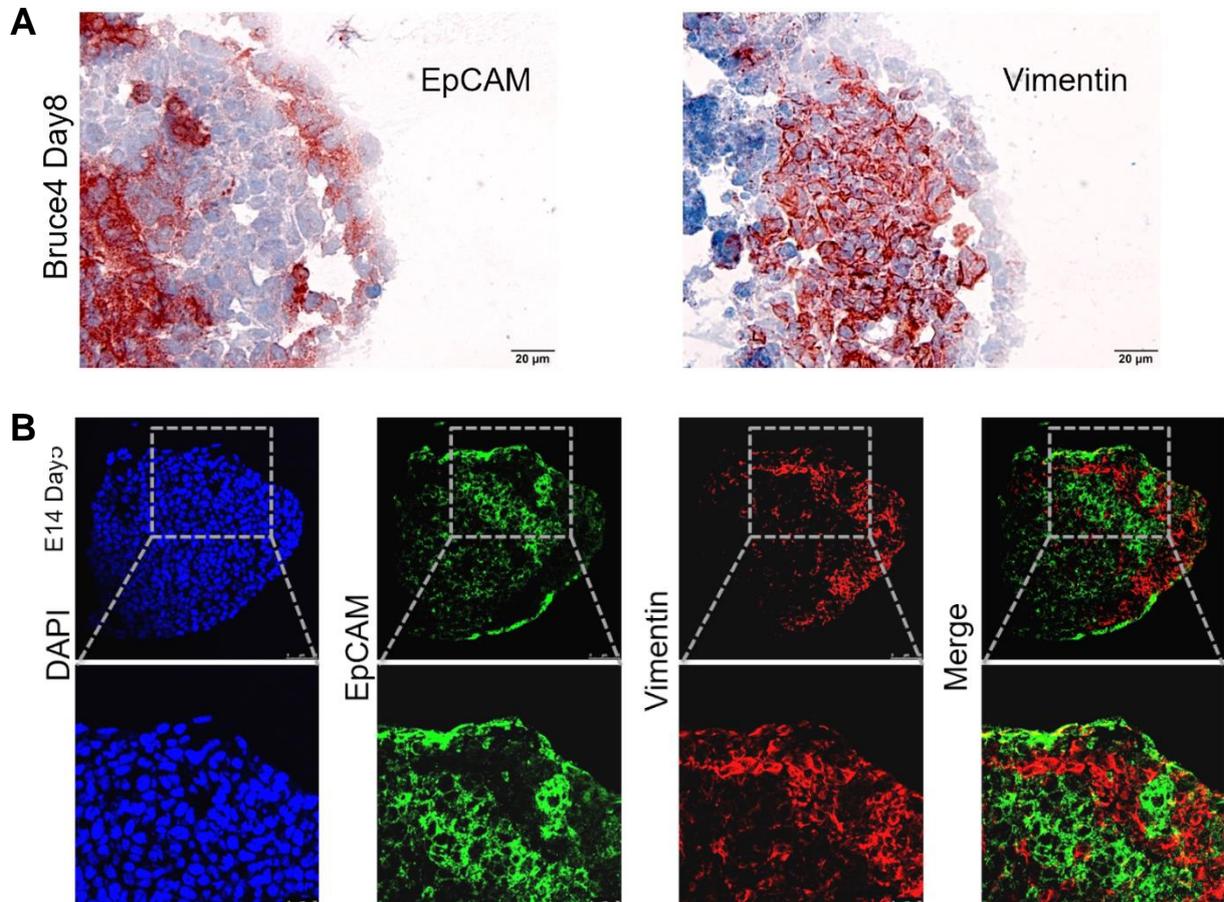


**Fig. 10: EpCAM and Foxa2 expression in differentiated ESCs**

Bruce4 and E14TG2 $\alpha$  ESCs were harvested, washed, and plated in hanging drops, and transferred after three days to ultra-low attachment plates. Subsequently, the EBs from day eight of Bruce4 ESCs and day five of E14TG2 $\alpha$  ESCs were shock-frozen, cryo-sectioned and were stained with EpCAM- and Foxa2-specific antibodies. (A) Shown are two representatives immunohistochemistry staining of EpCAM and Foxa2 in Bruce4 EBs sections (day eight) from  $n = 3$  independent experiments. (B) Shown are representative immunofluorescence double-staining of EpCAM and Foxa2 in E14TG2 $\alpha$  EBs (day five) with different magnifications. The images from stained slides of EBs were acquired via a laser scanning confocal microscope. Shown are the representative images from  $n = 3$  independent experiments with 6 EBs in each experiment.

#### 4.1.4.2 EpCAM expression and mesoderm

Vimentin is an intermediate filament protein that is highly expressed in mesenchymal cells and is frequently used as a marker for cells derived from the mesoderm, and for cells undergoing EMT in normal and malignant differentiation.



**Fig. 11: EpCAM and Vimentin expression in differentiated ESCs**

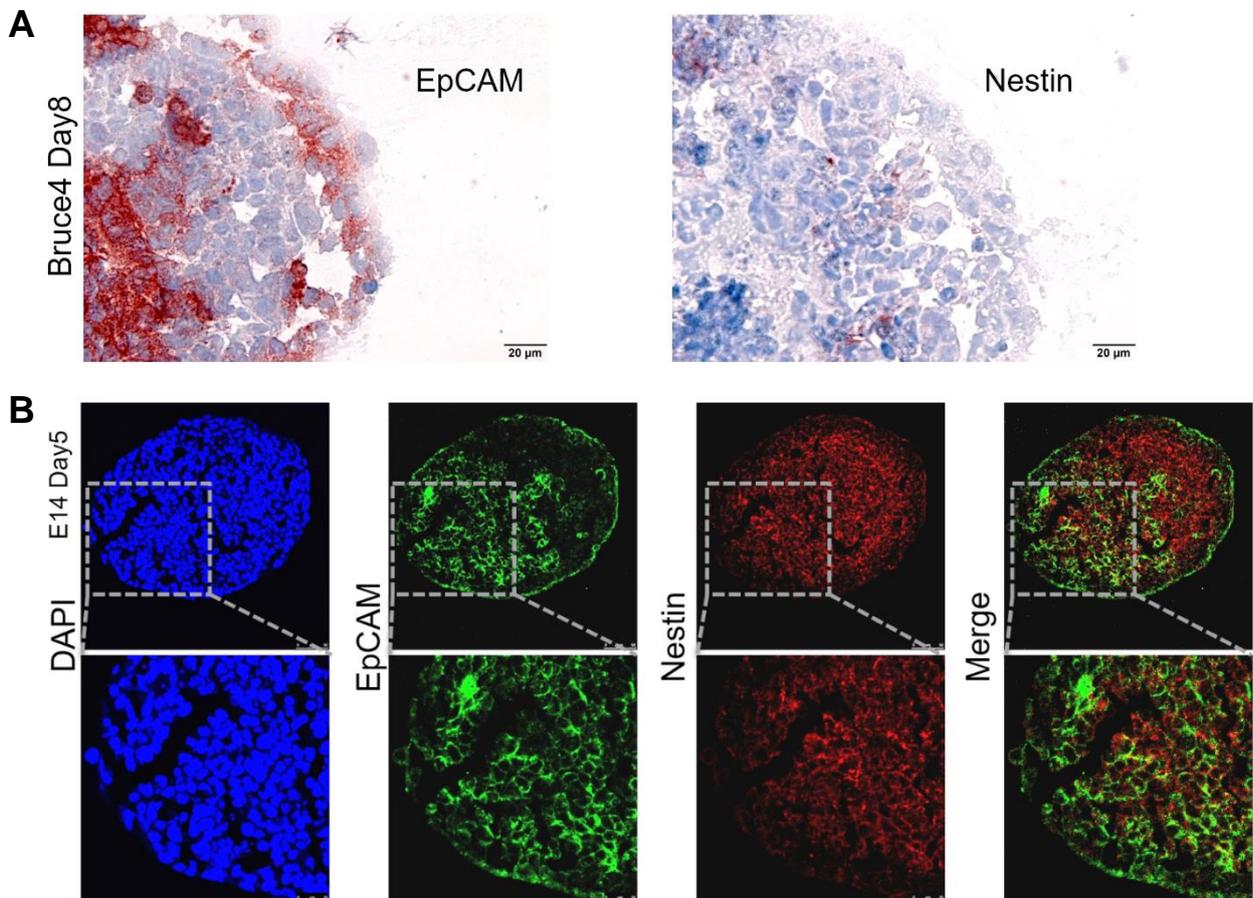
Bruce4 and E14TG2 $\alpha$  ESCs were harvested, washed, and plated in hanging drops, and transferred after three days to ultra-low attachment plates. Subsequently, the EBs from day eight of Bruce4 ESCs and day five of E14TG2 $\alpha$  ESCs were shock-frozen, cryo-sectioned, and were stained with EpCAM- and Vimentin-specific antibodies. (A) Shown are representative immunohistochemistry staining of EpCAM and Vimentin in Bruce4 EBs sections (day eight) from  $n = 3$  independent experiments. (B) Shown are immunofluorescence double-staining of EpCAM and Vimentin in E14TG2 $\alpha$  EBs (day five) with different magnifications. The images from stained slides of EBs were acquired via a laser scanning confocal microscope. Shown are the representative images from  $n = 3$  independent experiments with 6 EBs in each experiment.

**Figure 11A** shows consecutive sections of Bruce4 EBs stained with EpCAM- and Vimentin-specific antibody on day eight of spontaneous differentiation. A mutually exclusive

expression was found in a comparison of the staining of EpCAM and Vimentin in consecutive sections of EBs. To confirm this exclusive expression pattern, immunofluorescence double-staining of EpCAM and Vimentin was performed with E14TG2 $\alpha$  EBs on day five (**Figure 11B**). EpCAM<sup>+</sup> cells at the margin of and within EBs did not express Vimentin, while, conversely, EpCAM<sup>-</sup> cells revealed Vimentin<sup>+</sup>. Hence, 3D-differentiation of ESCs results in EpCAM<sup>+</sup>/Vimentin<sup>-</sup> and EpCAM<sup>-</sup>/Vimentin<sup>+</sup> cell cluster segregation in EBs.

#### 4.1.4.3 EpCAM and ectoderm

Nestin is an intermediate filament protein whose expression is widely used as a marker for stem cells in the developing nervous system and for *in vitro* cultured ectodermal cells. Here, Nestin served as an ectodermal marker to investigate a correlation of the expression of EpCAM and ectoderm.



**Fig. 12: EpCAM and Nestin expression in differentiated ESCs**

Bruce4 and E14TG2 $\alpha$  ESCs were harvested, washed, and plated in hanging drops, and transferred after three days to ultra-low attachment plates. Subsequently, the EBs from day eight of Bruce4 and day five were shock-frozen, cryo-sectioned and were stained with EpCAM- and Nestin-specific antibodies. (A) Shown are two representative immunohistochemistry staining of EpCAM and Nestin in Bruce4 EBs sections (day8) from n = 3 independent experiments. (B) Shown are immunofluorescence double-staining of EpCAM and Nestin in E14TG2 $\alpha$  EBs (day five) with different magnifications. The images from stained slides of EBs were acquired via a laser scanning confocal microscope. Shown are the representative images from n = 3 independent experiments with 6 EBs in each experiment.

**Figure12A** shows consecutive sections of Bruce4 EBs stained with EpCAM- and Nestin-specific antibodies on day eight of spontaneous differentiation in EBs. No obvious correlation of expression could be observed between EpCAM and Nestin in consecutive EBs sections. To confirm this, immunofluorescence double-staining of EpCAM and Nestin was performed with E14TG2 $\alpha$  EBs on day five of spontaneous differentiation in EBs. EpCAM expression was unrelated to Nestin, suggesting that the distribution of EpCAM is not correlated with ectodermal cells (**Figure 12B**).

Taken together, it can be concluded that EBs generated by the hanging drop method are an adequate 3D model to simulate early embryonic development *in vitro*, as EBs are comprised of all three germ layers. Secondly, EpCAM is generally down-regulated in the majority of differentiated cells within EBs following spontaneous differentiation. However, endodermal cells maintain the expression of EpCAM, especially within marginal cells representing Foxa2<sup>+</sup> visceral endoderm, whereas Vimentin<sup>+</sup> mesodermal cells are entirely devoid of EpCAM.

#### **4.1.5 Function of EpCAM in ESCs differentiation**

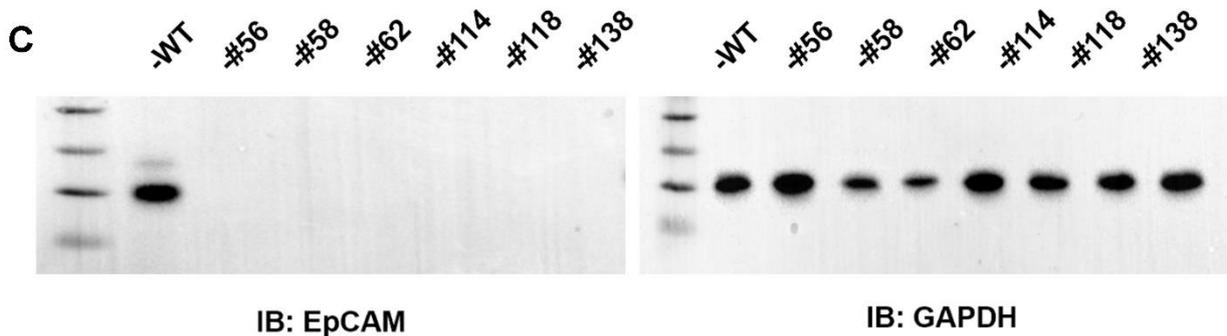
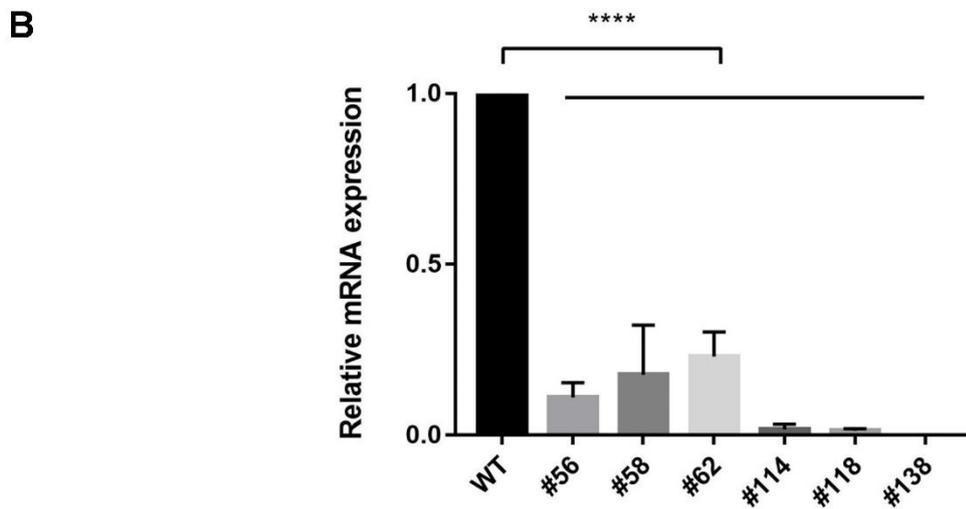
During early gastrulation, the expression of EpCAM demonstrated to be selective in space and time. Endodermal cells retained a strong expression of EpCAM, whereas EpCAM becomes entirely down-regulated in mesodermal progenitors. Based on our previous data, it was known that EpCAM over-expression inhibited mesodermal differentiation to

cardiomyocytes (Sarrach *et al.*, 2018), but that, however, mesodermal differentiation depended on EpCAM<sup>+</sup> endodermal cells. These results were achieved with the use of CRIPR-Cas 9-dependent knockout variants of EpCAM in E14TG2 $\alpha$  ESCs. A requirement of physical contact between EpCAM<sup>+</sup> Sox17-producing endodermal cells and mesodermal cells has been described in earlier publications (Holtzinger, Rosenfeld and Evans, 2010; Varner and Taber, 2012).

In this respect, the heart develops from mesoderm early in embryonic differentiation. Cardiac progenitors locate in the lateral plate mesoderm and maintain close contact with the underlying endoderm. Accordingly, co-culture of mouse ESCs with endodermal cell lines resulted in a strong induction of Flk1<sup>+</sup>/PDGFR $\alpha$ <sup>+</sup> cardiac progenitors in a dose-dependent fashion (Uosaki *et al.*, 2012). These results suggested that EpCAM<sup>+</sup> endodermal cells are mandatory for the formation of cardiomyocytes during development, however, the actual impact of a loss-of-function of EpCAM still remains largely unknown.

#### **4.1.5.1 Characterization of CRISPR/Cas9-mediated EpCAM knockout ESCs clones**

To understand embryonic development in the absence of EpCAM, knockout E14TG2 $\alpha$  clones were generated by CRISPR-Cas9-mediated genetic engineering with EpCAM-specific guide RNAs, and were confirmed *via* genomic DNA sequencing and protein expression. All EpCAM knockouts in ESCs used in experiments performed in the present thesis were homozygous clones.

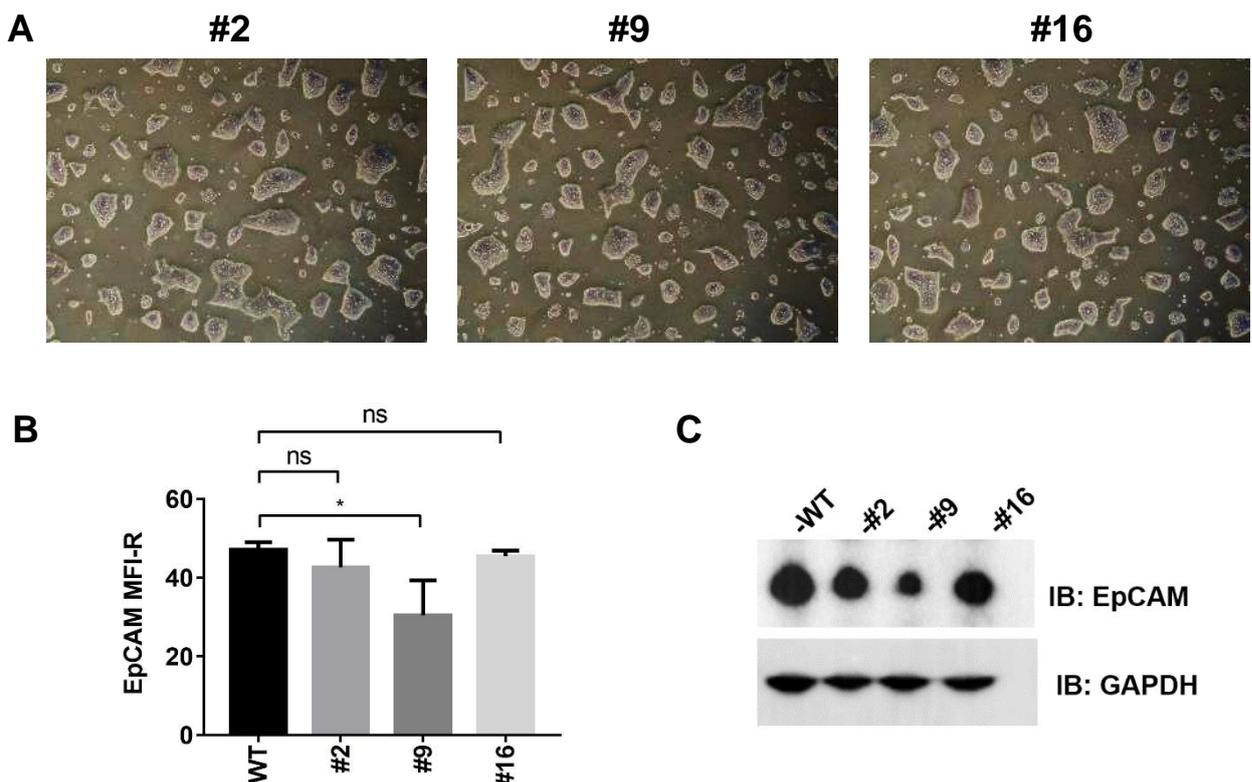


**Fig. 13: Characterization of EpCAM knockout E14TG2 $\alpha$  ESCs clones**

E14TG2 $\alpha$  wildtype (WT) and EpCAM single knockout clones #56, #58, #62, #114, #118, and #138 ESCs were cultured in gelatin-coated flasks, and images were acquired every week. (A) Shown are EpCAM knockout clones #56, #58, #62, #114, #118, and #138 ESCs maintained under pluripotency conditions in the presence of LIF. (B) E14TG2 $\alpha$  WT and EpCAM knockout clones ESCs were collected, RNA was isolated, and EpCAM mRNA levels were assessed via qPCR. Shown are the mean values

with standard deviations from  $n = 3$  independent experiments. p-value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \*\*\*\* $<0.0001$ . (C) Whole cell extracts of E14TG2 $\alpha$  WT and EpCAM knockout clones ESCs were generated, proteins were separated via SDS-PAGE and detected with EpCAM-specific antibody. Staining of GAPDH served as loading control for equal protein amounts. Shown are representative immunoblot results from  $n = 3$  independent experiments.

E14TG2 $\alpha$  WT and EpCAM knockout clones (#56, #58, #62, #114, #118, #138) were used in this study. All cell lines were cultured under pluripotent conditions in the presence of LIF. Images of the morphology of each ESC clone were acquired weekly, and their typical round-shaped colonies were visualized *via* microscopy. **Figure 13A** shows a typical ESC morphology of all knockout clones (#56, #58, #62, #114, #118, #138). EpCAM mRNA and protein levels were assessed in WT and EpCAM-knockout ESC clones using quantitative RT-PCR and immunoblot staining. EpCAM knockout clones displayed severally reduced EpCAM mRNA levels (**Figure 13B, C**). Residual mRNA levels of *EPCAM* following CRISPR/Cas9 knockout may result from non-homologous end-joining of double-strand breaks. Despite these residual mRNA levels, protein expression was entirely lost in all  $n = 6$  EpCAM knockout ESCs clones.

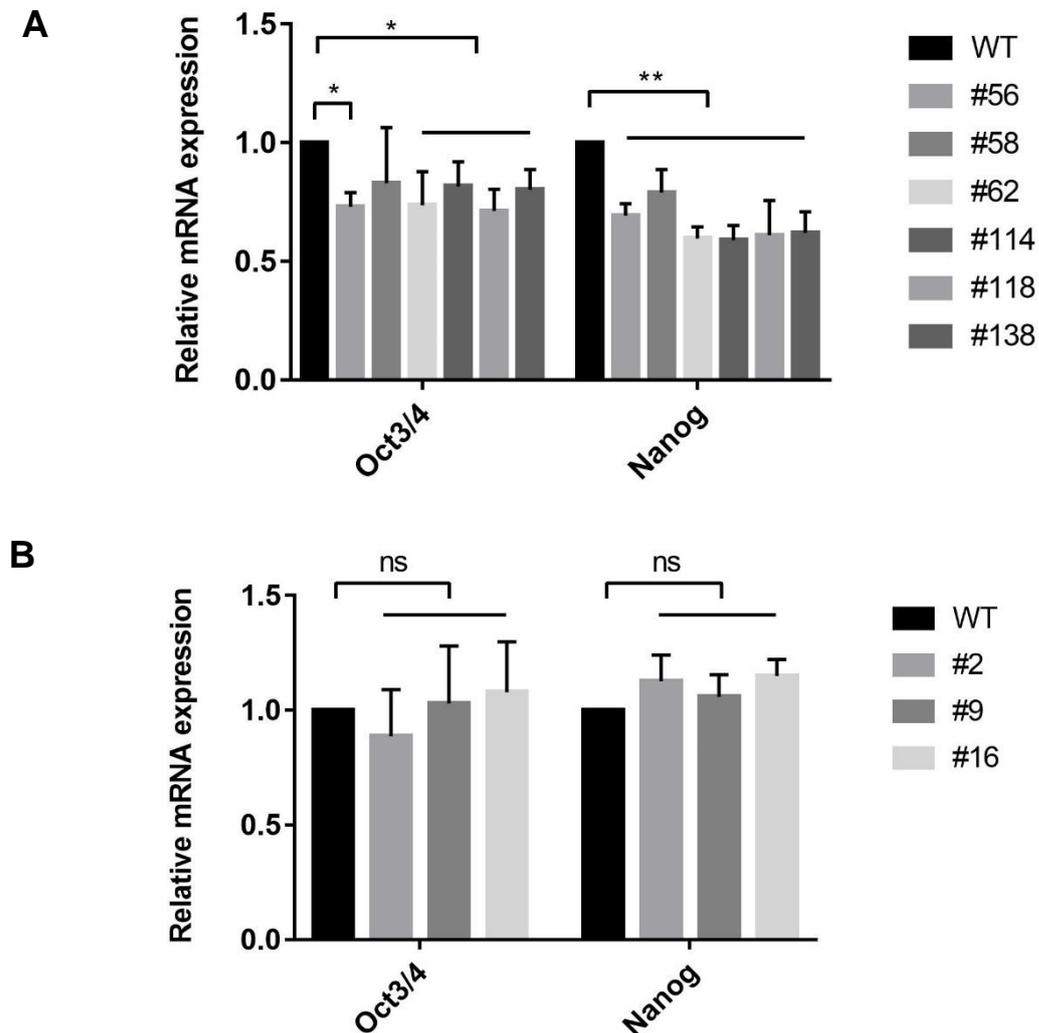


**Fig. 14: Characterization of CRISPR/Cas9 control E14TG2 $\alpha$  ESCs clones**

E14TG2 $\alpha$  WT and control clones #2, #9, #16 were compared. CRISPR/Cas9 knockout control clones #2, #9, and #16 represent single cell clones of E14TG2 $\alpha$  ESCs that have been transfected with guide RNA and Cas9, but did not show any obvious differences in EpCAM expression compared to WT E14TG2 $\alpha$  ESCs. All cell lines were plated in gelatin-coated flasks and images were acquired every week. (A) Shown are  $n = 3$  CRISPR/Cas9 control clones that were regularly maintained under pluripotency conditions in the presence of LIF. Images are shown from representative colonies. (B) Whole cell extracts of E14TG2 $\alpha$  WT and control clones #2, #9, and #16 ESCs were collected, proteins were separated in SDS-PAGE, and EpCAM was detected with specific antibody. Staining of GAPDH served as loading for equal protein amounts. Shown are representative immunoblot results from  $n = 3$  independent experiments. (C) E14TG2 $\alpha$  WT and control clones (#2, #9, #16) ESCs were collected, washed and determined by flow cytometry using EpCAM-specific antibody. Shown are mean fluorescence intensity ratio (MFI-R) with standard deviations from  $n = 3$  independent experiments.  $p$ -value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \* $<0.05$ ; ns $>0.05$ .

Apart from  $n = 6$  EpCAM knockout clones, three additional clones (#2, #9, #16) were selected as controls. These controls were generated through transfection of the CRISPR/Cas9 all-in-one plasmid containing an expression cassette for Cas9, an EpCAM-specific guide-RNA, and GFP as a marker for FACS-based selection of transfected cells. Single cell clones were generated and three clones, which did not show any apparent differences in EpCAM expression compared to WT E14TG2 $\alpha$  ESCs, were further analyzed. The morphology of each cell clones was recorded under pluripotency conditions in the presence of LIF in the culture medium. All three CRISPR/Cas9 control clones showed a typical stem cell colony morphology (**Figure 14A**). The presence of EpCAM protein at the cell surface and in cell lysates was assessed by flow cytometry and immunoblot, respectively. Flow cytometry results showed single cell E14TG2 $\alpha$  clone #9 has 40% less EpCAM expression compared to WT E14TG2 $\alpha$  with statistical difference (**Figure 14B**), while the two remaining control clones exhibit a comparable expression level of EpCAM. Similar to flow cytometry data, immunoblot results displayed a clear expression of EpCAM protein at 35 kDa, which was reduced in single cell E14TG2 $\alpha$  clone #9 (**Figure**

14C). The results suggested that two out of three control clones expressed comparable level of EpCAM of WT E14TG2 $\alpha$  ESCs, while clone #9 displayed approx. halved levels of



EpCAM.

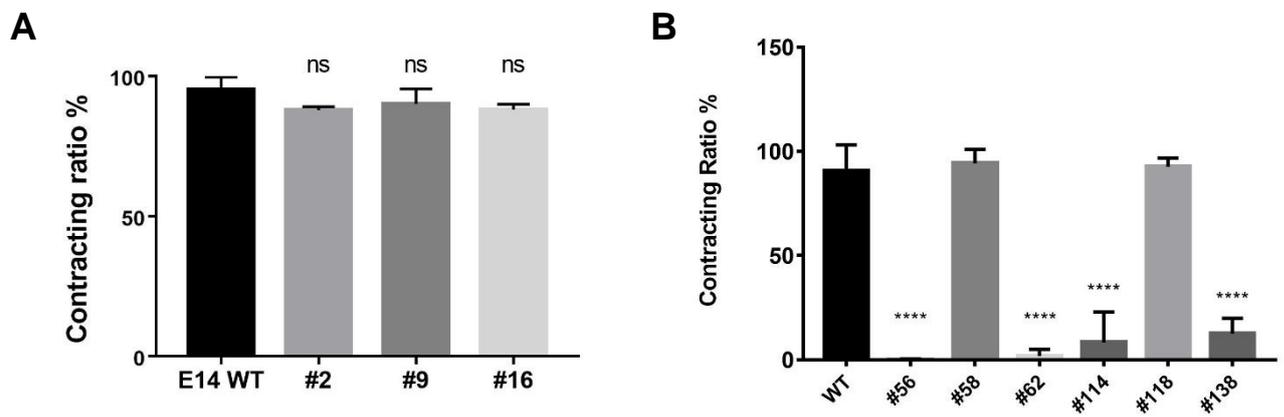
**Fig. 15: Expression of the pluripotency markers Oct3/4 and Nanog in WT E14TG2 $\alpha$  and CRISPR/Cas9 control clones**

E14TG2 $\alpha$  WT, CRISPR/Ca9 EpCAM-knockout clones #56, #58, #62, #114, #118, #138 and CRISPR/Cas9 control clones #2, #9, #16 were plated in gelatin-coated flask under pluripotency conditions in the presence LIF. RNA was isolated from cells and qPCR was performed to detect Oct3/4 and Nanog mRNA expression levels as pluripotency markers. (A) Shown are E14TG2 $\alpha$  WT and EpCAM knockout clones. (B) Shown are E14TG2 $\alpha$  WT and CRISPR/Cas9 control clones. The expression levels of Oct3/4 and Nanog mRNA were normalized for the expression in E14TG2 $\alpha$  WT cells. Results are presented as means with standard deviations from n = 3 independent experiments. p-value were calculated with a Two-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \*<0.05; \*\*<0.01; ns>0.05.

Next, the pluripotency of E14TG2 $\alpha$  WT and knockout clones was evaluated in all single cell clones under pluripotency conditions, *i.e.* in the presence of LIF in the culture media. E14TG2 $\alpha$  EpCAM knockout clones were characterized by reduced expression of pluripotent genes Oct3/4 and Nanog, ranging from 20%–48% and 57%–75% reduction, respectively (**Figure 15A**), confirming that EpCAM was involved in the maintenance of stem cell pluripotency. Three CRISPR/Cas9 control clones displayed levels of Oct3/4 and Nanog mRNA comparable to E14TG2 $\alpha$  WT cells, thus suggesting full pluripotency of these control clones.

#### 4.1.5.2 Differentiation of E14TG2 $\alpha$ WT cells and CRISPR/Cas9 derivatives

Upon spontaneous differentiation in EBs, E14TG2 $\alpha$  WT cells generate contracting cardiomyocytes *in vitro*. Thus, the differentiation of E14TG2 $\alpha$  WT cells and CRISPR/Cas9 derivatives was monitored regarding the formation of cardiomyocytes (EBs contracting ratio) and the size of EBs. None of the EpCAM<sup>+</sup> CRISPR/Cas9 control clones was impaired in cardiomyocyte development, as measured through the rates of contracting EBs and compared to E14TG2 $\alpha$  WT cells (**Figure 16A**). In contrast, four out of six EpCAM knockout clones were severely impaired in the formation of contracting EBs, with contraction rates dropping to 0.1–12.5% (**Figure 16B**). These data demonstrated that the genetic knockout of EpCAM had a negative impact on the spontaneous differentiation of E14TG2 $\alpha$  WT cells into contracting cardiomyocytes in EBs.



**Fig. 16: Differentiation of E14TG2 $\alpha$  WT cells and CRISPR/Cas9 derivatives into contracting cardiomyocytes**

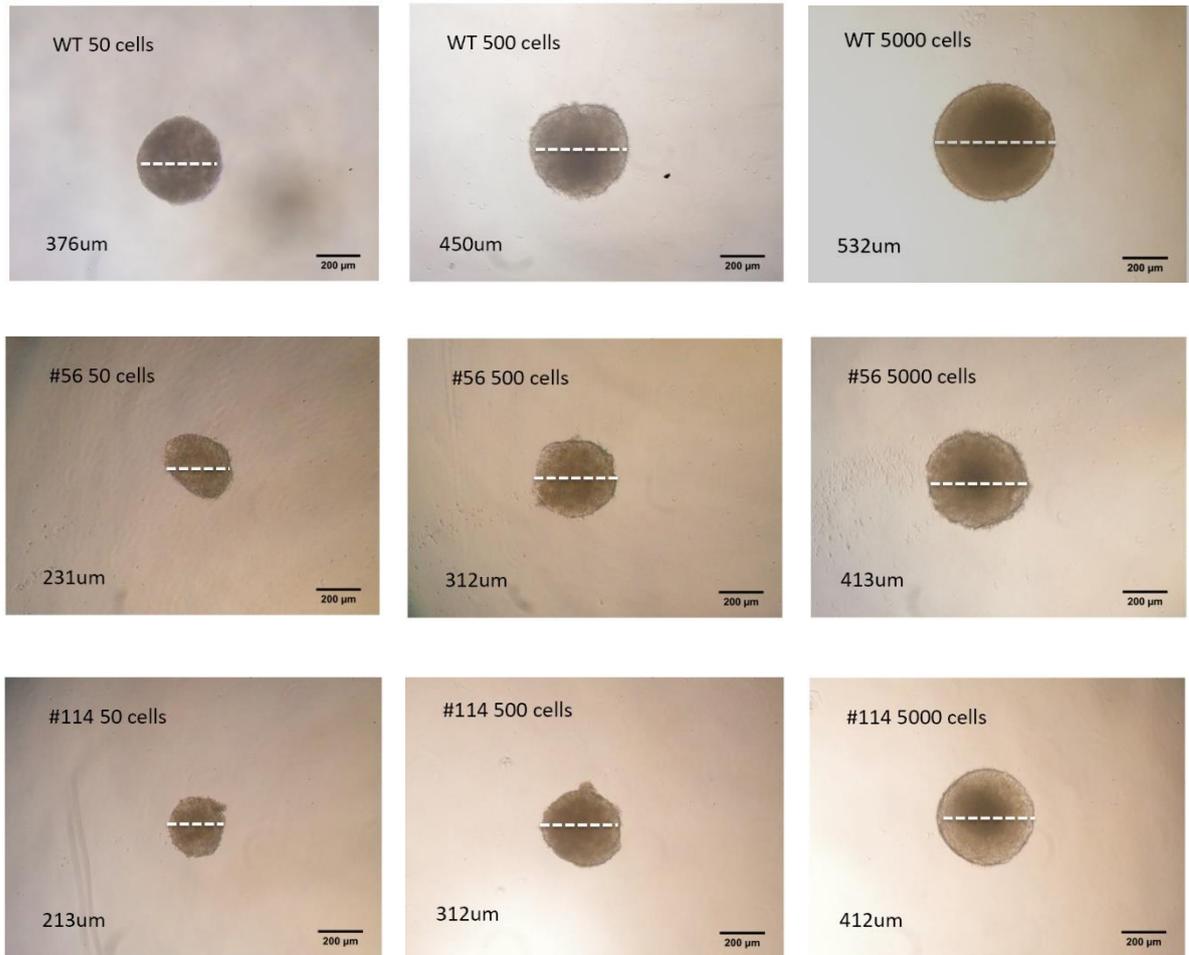
E14TG2 $\alpha$  WT and CRISPR/Cas9 EpCAM knockout cell lines were harvested, washed, and plated in hanging drops, and transferred after three days to ultra-low attachment plates. On day seven, EBs were transferred to 96-well plate with fresh medium. On day 10 of the differentiation of EBs, the percentage of contracting EBs was recorded and normalized to the numbers of total EBs. (A) Shown are the mean percentages with standard deviations of contracting EBs of E14TG2 $\alpha$  WT and CRISPR/Cas9 control clones from  $n = 3$  independent experiments with  $n \geq 80$  EBs in each experiment. (B) Shown are the mean percentages with standard deviations of contracting EBs of E14TG2 $\alpha$  WT and CRISPR/Cas9 EpCAM knockout cell lines from  $n = 3$  independent experiments with  $n \geq 80$  EBs in each experiment. p-value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \*\*\*\* $<0.0001$ ; ns $>0.05$ .

#### 4.1.5.3 Impact of EB size on spontaneous differentiation of ESCs

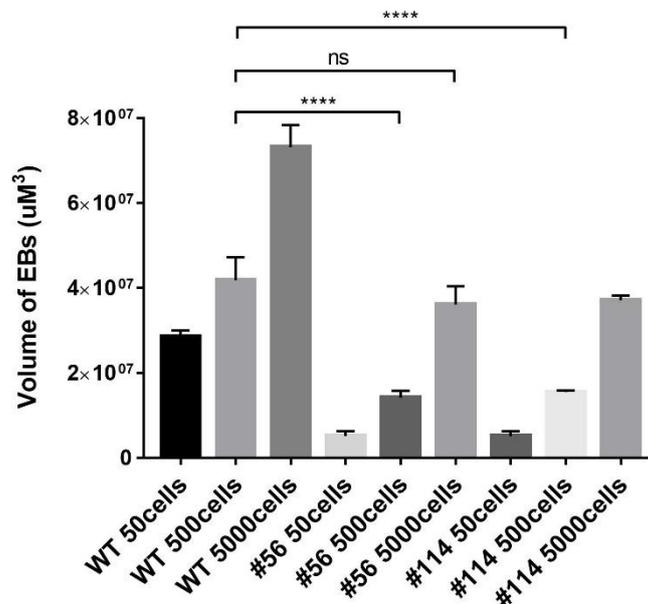
During the generation of EBs, E14TG2 $\alpha$  WT and EpCAM knockout clones were characterized by differences in EBs size. EpCAM knockout clones revealed significantly smaller than WT EBs at day five of differentiation (**Figure 17A**). It has been reported that the size of EBs plays a role in the efficiency of cardiomyocyte formation (Hwang *et al.*, 2009). Therefore, the potential impact of differing EBs sizes on the formation of

cardiomyocytes was evaluated with E14TG2 $\alpha$  WT and EpCAM knockout clones.

**A**



**B**

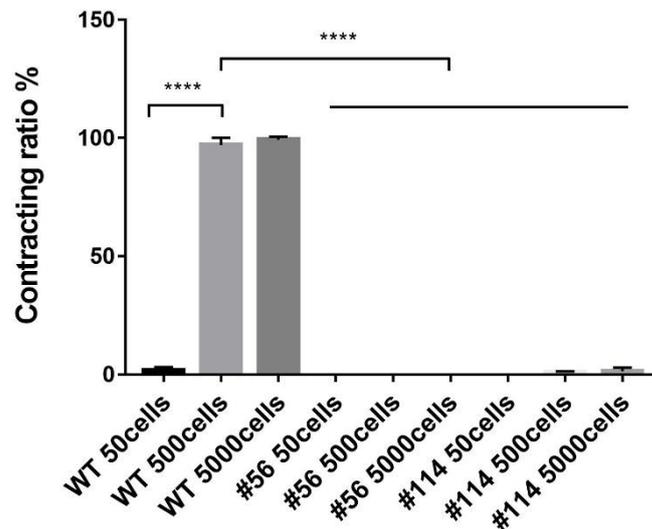


**Fig. 17: Correlation of ESCs numbers with EBs size *in vitro***

E14TG2 $\alpha$  WT ESCs and CRISPR/Cas9 EpCAM knockout clones #56, #114 were harvested, washed, and plated in hanging drops with different initial number of 50, 500 and 5,000 cells, respectively. EBs were transferred to ultra-low attachment plates after three days. On day seven, EBs were transferred

again to 96-well plate with fresh medium. (A) On day five of the differentiation of ESCs in EBs, images of representative EBs were acquired *via* microscopy. (B) From these microscopic images, diameters were measured and volumes of EBs were extrapolated. Shown are mean volumes with standard deviations from n = 3 independent experiments with n = 27 EBs in each experiment. p-value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \*\*\*\*<0.0001; ns>0.05.

Different initial number of cells of E14TG2 $\alpha$  WT and EpCAM knockout clones #56 and #114 (50, 500, and 5,000) were plated in 96-well plates to generate EBs. Representative pictures of EBs at day five including diameters are depicted in **Figure 17A**. Volume extrapolation of EBs was performed with three independent biological repeats, and mean and SD are presented in **Figure 17B**. Generally, E14TG2 $\alpha$  WT ESCs generated EBs of bigger size compared with both CRISPR/Cas9 EpCAM knockout clones at equal starting cell numbers (**Figure 17B**). A tenfold excess of CRISPR/Cas9 EpCAM knockout clones #56 and #114 (5,000 cells) was required to generate EBs of equal size to E14TG2 $\alpha$  WT cells (500 cells) (**Figure 17B**), suggesting that the knockout of EpCAM in ESCs impacted on the capacity of cells to proliferate during differentiation in EBs.



**Fig. 18: Correlation of EBs size with cardiomyocyte formation *in vitro***

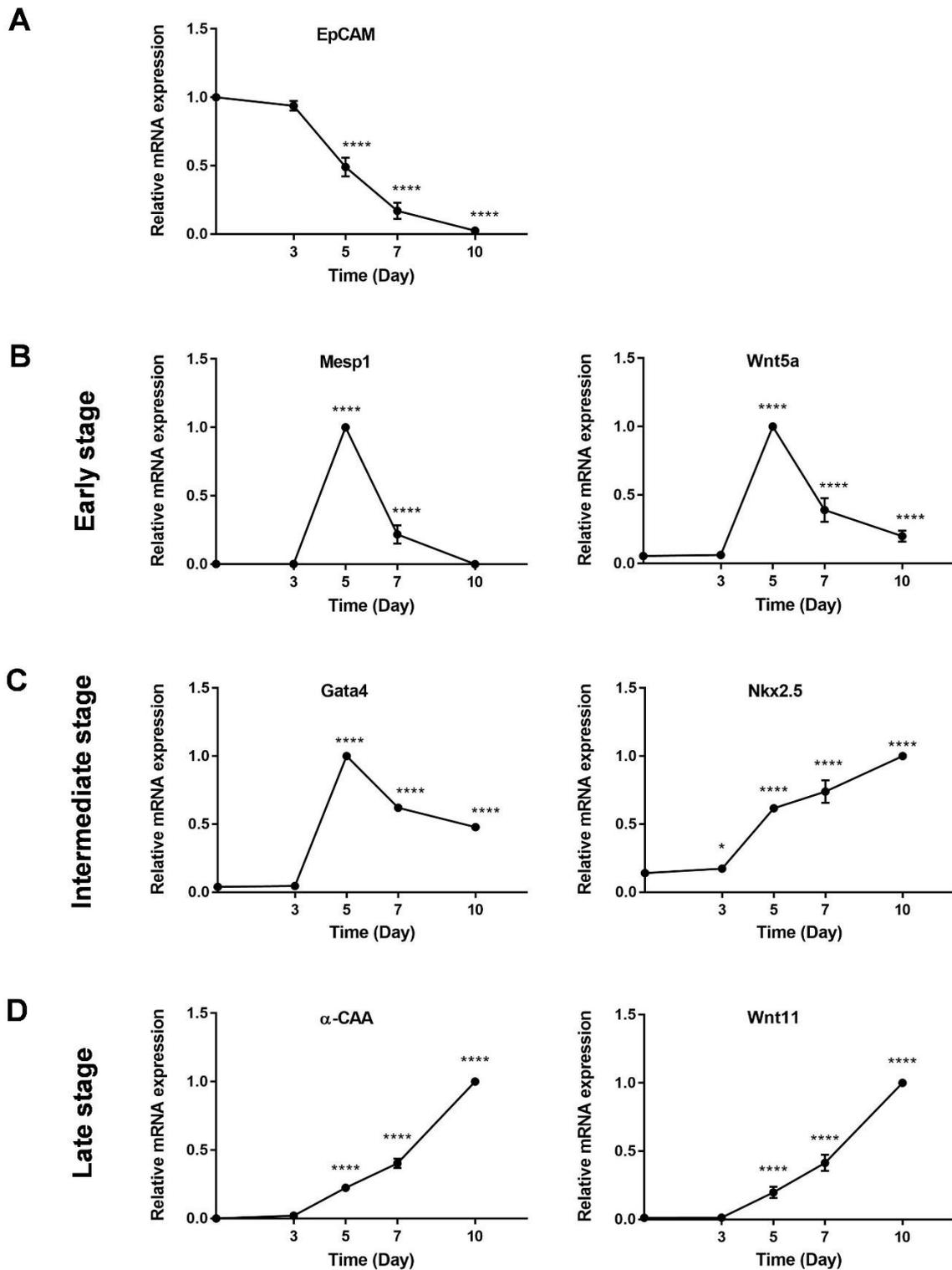
E14TG2 $\alpha$  WT ESCs and CRISPR/Cas9 EpCAM knockout clones #56, #114 were harvested, washed, and plated in hanging drops with different initial cell numbers of 50, 500, and 5,000 cells, respectively. EBs were transferred after three days to ultra-low attachment plates. On day seven, EBs were transferred again to 96-well plate with fresh medium. On day 10 of the differentiation of EBs, the

percentage of contracting EBs were recorded and normalized to the numbers of total EBs. Shown are the mean percentages with standard deviations of contracting EBs from  $n = 3$  independent experiments with  $n = 27$  EBs in each experiment. p-value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \*\*\*\* $<0.0001$ .

In a following experiment, the contraction rate of EBs from E14TG2 $\alpha$  WT and EpCAM knockout clones #56 and #114 with starting cell numbers of 50, 500, and 5,000 cells was assessed in a 10 days differentiation. The aim of this experiment was to assess whether equivalent EBs size can rescue the cardiomyocyte formation defect of EpCAM knockout clones. Although EBs generated from CRISPR/Cas9 EpCAM knockout clones could reach a size equivalent to E14TG2 $\alpha$  WT EBs at a tenfold cell excess, neither EpCAM knockout clone #56 nor #114 was able to generate functional cardiomyocytes (**Figure 18**). While E14TG2 $\alpha$  WT EBs required a starting cell number of 500 to display high percentages of contracting EBs, EpCAM knockout clones #56 and #114 were incapable of differentiating to contracting cardiomyocytes, even at starting seeding number of 5,000 cells (**Figure 18**).

#### **4.1.5.4 Mesp1 is important for cardiomyocyte differentiation**

Cardiogenesis is a complex process that is organized and orchestrated by multiple genes. Mesoderm posterior protein 1 (Mesp1) acts as a master regulator in cardiomyocyte formation. Genome-wide measurements of RNA revealed Mesp1 is able to activate and inhibit cardiac-associated genes. Chromatin immunoprecipitation data proved Mesp1 can directly bind to the promoter of many key genes associated with the cardiac transcriptional machinery, to upregulate their expression (Bondue *et al.*, 2008). Throughout cardiomyocyte formation, mesodermal progenitors initially require a Mesp1/Wnt5a-dependent activation, which is then followed by a reduction of Wnt5a and Mesp1, and finally the induction of Wnt11 expression in order to complete cardiomyocyte maturation via the physical contact and instruction with Sox17<sup>+</sup>/EpCAM<sup>+</sup> endodermal cells (Hwang *et al.*, 2009; Holtzinger, Rosenfeld and Evans, 2010; Mazzotta *et al.*, 2016a).

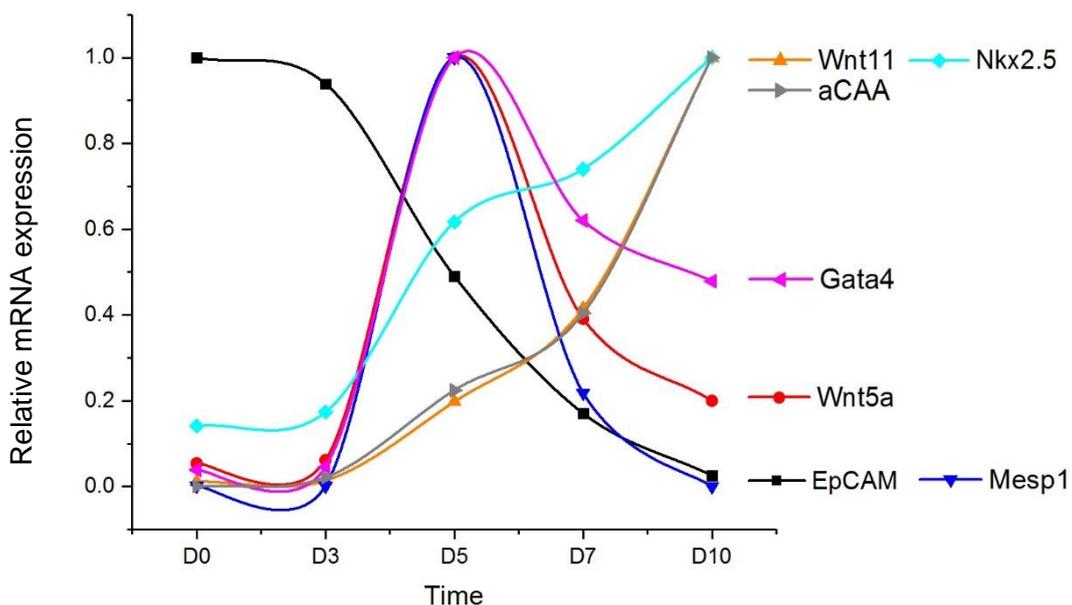


**Fig. 19: Analysis of the transcriptional expression dynamics of cardiomyocyte markers during E14TG2 $\alpha$  differentiation**

E14TG2 $\alpha$  WT ESCs were harvested, washed, and plated in hanging drops, and differentiated for 10 days. EBs were collected at each time points of day 0, 3, 5, 7, 10, RNA of EBs was isolated, and EpCAM (A), Mesp1, Wnt5a (B), Gata4, Nkx2.5 (C) and  $\alpha$ -CAA, Wnt11 (D) expression on transcriptional level were measured by qRT-PCR. Values at different time points are normalized to the highest value in each group. Shown are the means and standard deviations of EpCAM and cardiomyocyte markers

expression during EBs differentiation from  $n = 3$  independent experiments. p-value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \* $<0.05$ ; \*\*\*\* $<0.0001$ .

Spontaneous differentiation of E14TG2 $\alpha$  WT cells was performed in a kinetic experiment over 10 days, and expression of EpCAM and various genes was measured by qRT-PCR. Here, Wnt5a, Mesp1 as early regulators, Gata4, Nkx2.5 as intermediate regulators, Wnt11 as a late regulator, and  $\alpha$ -CAA as a marker for matured cardiomyocyte were assessed. The assessment of mRNA levels of these genes showed a time dependency, with a peak of Wnt5a, Mesp1 (**Figure 19B**), and Gata4 mRNA expression at day five (**Figure 19C**), and a strong or complete loss of Wnt5a and Mesp1 (**Figure 19B**) at day seven, respectively. Gata4 was decreased to 48% at day 10 (**Figure 19C**). Starting from day five, Wnt11, Nkx2.5, and  $\alpha$ -CAA expression was gradually increased and peaked at day 10 (**Figure 19C, D**).



**Fig. 20: Kinetic expression of the mRNA of cardiogenesis-associated genes during E14TG2 $\alpha$  WT ESCs differentiation in EBs**

Summary of the results of the kinetics of cardiogenesis-associated genes. E14TG2 $\alpha$  WT ESCs were harvested, washed, and plated in hanging drops, and differentiated for 10 days. EBs were collected at each time points of day 0, 3, 5, 7, 10, RNA of EBs was isolated, EpCAM, Mesp1, Wnt5a, Gata4, Nkx2.5,  $\alpha$ -CAA, and Wnt11 expression at the transcriptional level were measured by qRT-PCR. Shown are the mean values of the expression of EpCAM and cardiomyocyte markers at the transcriptional level during

EBs differentiation from n = 3 independent experiments. Values at different time points are normalized to the highest value in each group. EpCAM, Mesp1, Wnt5a, gata4, nkx2.5, Wnt11 and Nkx2.5 are presented as black, blue, red, purple, cyan, orange and grey curve, respectively.

Taken together, Mesp1 appears to be required as a master regulator to induce Gata4 and Nkx2.5 expression at early stage of ESCs differentiation into cardiomyocytes. At later stages, the expression of Mesp1 and wnt5a has to be down-regulated, and expression of wnt11 induced, to allow differentiating ESCs to fully mature  $\alpha$ -CAA<sup>+</sup> cardiomyocytes (Figure 20).

#### 4.1.5.5 EpCAM knockout impacts on cardiomyocyte differentiation

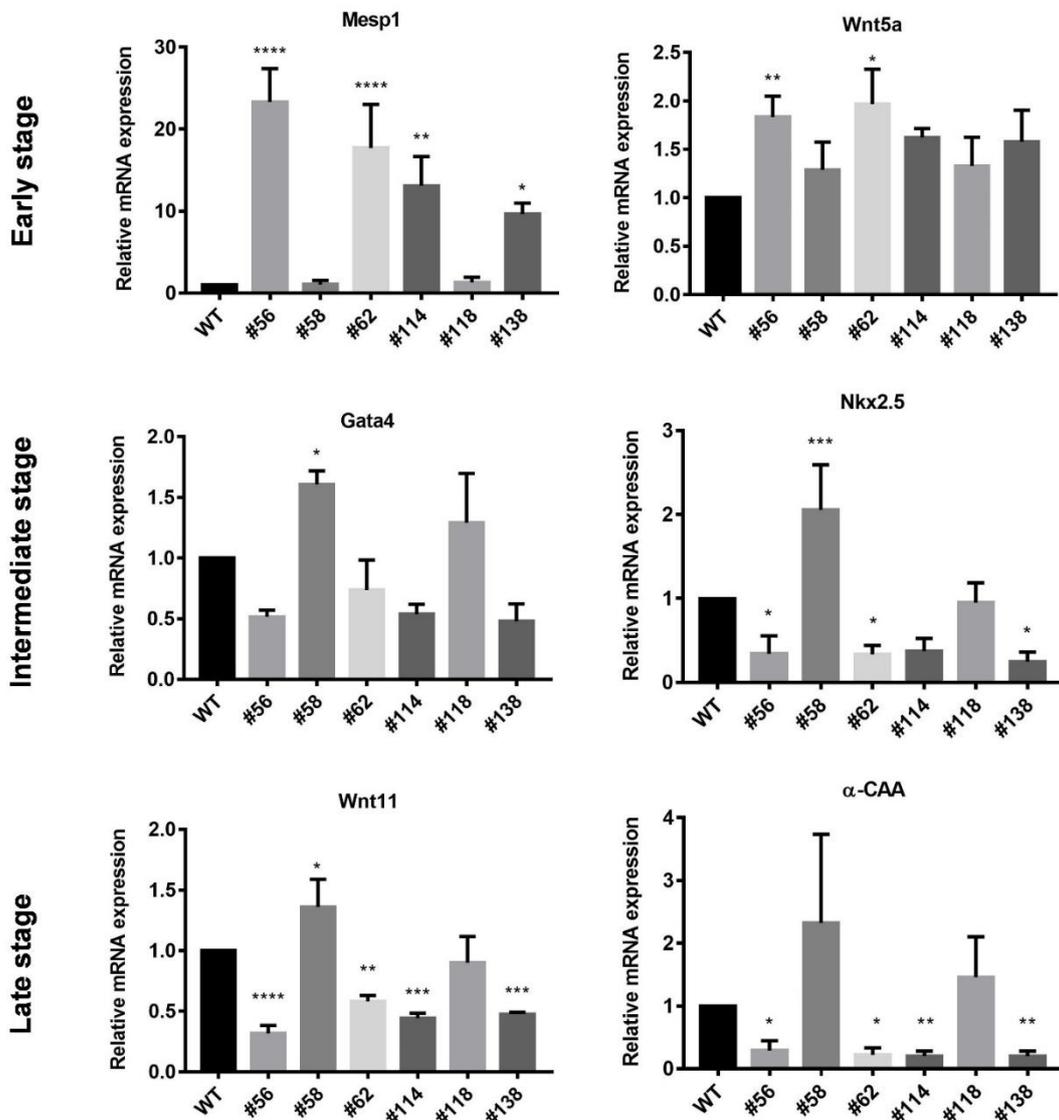


Fig. 21: EpCAM knockout impacts on the expression of genes related to cardiomyocyte differentiation

E14TG2 $\alpha$  WT ESCs and CRISPR/Cas9 EpCAM knockout clones #56, #58, #62, #114, #118, and #138 were harvested, washed, and plated in hanging drops, and transferred after three days to ultra-low attachment plates. On day seven, EBs were transferred to 96-well plate with fresh medium. RNA was isolated from EBs on day 10, and the expression of the mRNA of the cardiomyocyte differentiation-associated genes *Mesp1*, *Wnt5a*, *Gata4*, *Nkx2.5*, *Wnt11*, and  $\alpha$ -CAA was assessed by qRT-PCR. Shown are the means and standard deviations as relative mRNA expression levels of  $n = 3$  independent experiments with  $n = 27$  EBs in each experiment. All values were normalized to the value of E14TG2 $\alpha$  WT ESCs in each experiment. p-value were calculated with a One-way ANOVA test with multiple *posthoc* testing and Bonferroni correction. \* $<0.05$ ; \*\* $<0.01$ ; \*\*\* $<0.001$ ; \*\*\*\* $<0.0001$ .

To understand the dysfunction of cardiomyocyte formation in non-contraction EBs derived from CRISPR/Cas9 EpCAM knockout clones, the abovementioned cardiomyocytes genes were assessed at day 10 in E14TG2 $\alpha$  WT and CRISPR/Cas9 EpCAM knockout clones #56, #58, #62, #114, #118, and #138.

Unlike E14TG2 $\alpha$  WT and EpCAM knockout clones that generated contracting EBs (*i.e.* #58 and #118), EpCAM knockout clones #56, #62, #114, and #138 did not efficiently downregulate *Mesp1* expression, and eventually expressed 10- to 20-fold higher *Mesp1* mRNA levels compared to E14TG2 $\alpha$  WT on day 10 (**Figure 21**). Accordingly, slightly increased levels of *Wnt5a* were measured in these non-contracting E14TG2 $\alpha$  clones, too, and genes associated with cardiomyocyte differentiation including *Gata4*, *Nkx2.5*, *Wnt11*, and  $\alpha$ -CAA were significantly reduced compared to WT EBs at day 10 of spontaneous differentiation (**Figure 21**).

Taken together, the differentiation of EBs of non-contracting E14TG2 $\alpha$  EpCAM knockout clones was blocked at a *Mesp1*<sup>+</sup> stage. As depicted in **Figure 20**, *Mesp1* displays a biphasic expression pattern characterized by the induction of its expression around day three of differentiation with a peak of expression at day five. Throughout further differentiation, *Mesp1* expression is down-regulated, while *Wnt11*, *Nkx2.5*, and  $\alpha$ -CAA become up-regulate at day 10. However, non-contracting clones maintained high levels of expression of *Mesp1* and, though much less pronounced, of *wnt5a* expression at day 10.

These data indicate that the differentiation of EpCAM knockout EBs was blocked at a *Mesp1*<sup>+</sup> stage (**Figure 20**).

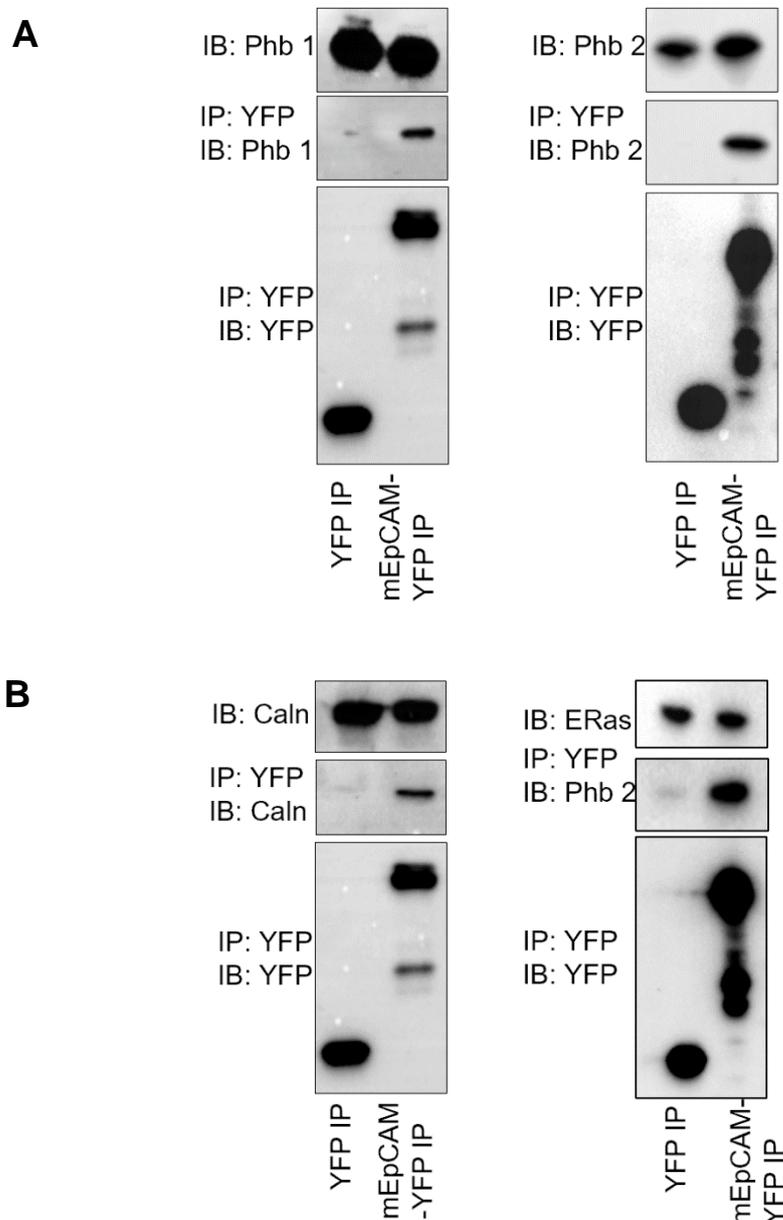
#### **4.1.6 EpCAM regulates ESCs differentiation via the ERas/AKT cascade**

From previous data obtained in our research group, it was known that EpCAM cleavage products EpCTF and EpICD, which are generated through regulated intramembrane proteolysis, do not inhibit cardiomyocyte formation. Constitutive expression of either EpCTF or EpICD in E14TG2 $\alpha$  did not impair cardiomyocyte formation, suggesting an inhibitory function for full-length EpCAM. Furthermore, although a knockout of the *EPCAM* gene impaired the cardiomyocyte formation, two out of six clones remained capable of generating functional cardiomyocyte. In order to understand the mechanism behind these somewhat conflicting results, potential binding partners and down-stream signaling of EpCAM were further investigated.

##### **4.1.6.1 Characterization of ERas as an EpCAM-binding partner**

To assess the interacting partners of full-length EpCAM, stable isotope labeling with amino acids in cell culture (SILAC), immunoprecipitation and liquid chromatography–mass spectrometry (LC-MS/MS) were performed by Matthias Hachmeister (Ph.D. student in our group). A total of 77 candidates for EpCAM-interaction partners were identified through a combination of SILAC and LC-MS/MS with the mouse teratocarcinoma cell line mF9. Prohibitin 1/2 represented top ranking interaction candidates, while Calnexin and ERas represented potential candidates with inferior enrichment scores.

Next, interactions of EpCAM with prohibitin 1 and 2, calnexin, and ERas were validated in lysates of E14TG2 $\alpha$  cells expressing either EpCAM fused to the yellow fluorescent protein (EpCAM-YFP) or YFP as a control.



**Fig. 22: Characterization of the interaction of EpCAM with Prohibitin 1 and 2, Calnexin, and ERas in E14TG2 $\alpha$  ESCs**

E14TG2 $\alpha$  ESCs stably expressing EpCAM-YFP or YFP were cultured under pluripotency conditions in the presence of LIF, cells were harvested, and cell lysates were used for co-immunoprecipitation experiments. EpCAM-YFP and YFP proteins were enriched with GFP-trap beads and were separated on SDS-PAGE. The co-precipitation of (A) prohibitin 1 and 2, and (B) calnexin and ERas were assessed with specific antibodies upon immunoblotting. Equal expression of prohibitin 1 and 2, calnexin, and ERas in both E14TG2 $\alpha$  derivative cell lines were confirmed by immunoblotting of whole cell lysates. Comparable immunoprecipitation of EpCAM-YFP and YFP was assessed in immunoblotting with GFP-

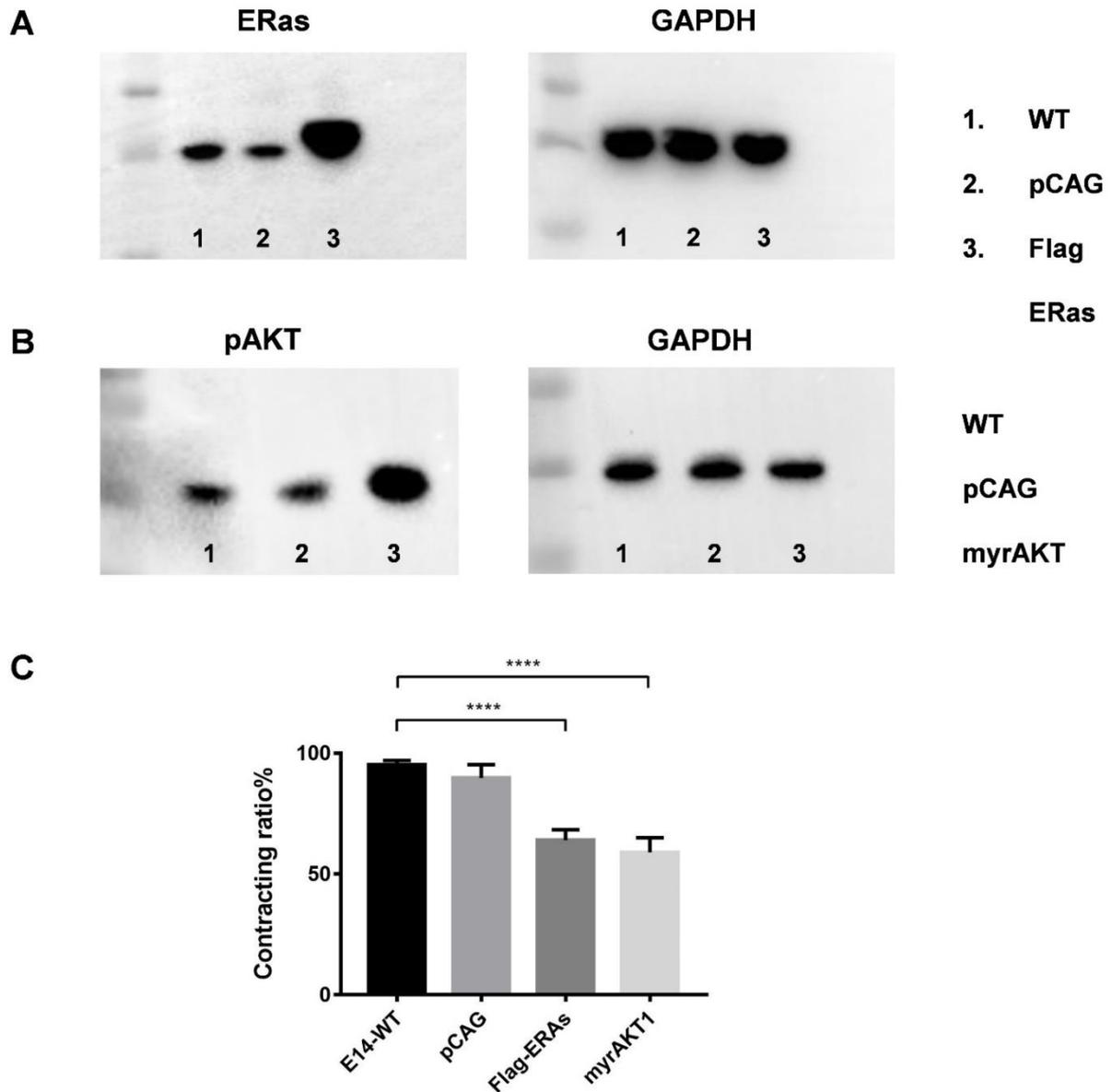
specific antibody (recognizing YFP). Shown are representative immunoblots from  $n = 3$  independent experiments.

**Figure 22** shows co-immunoprecipitation results of prohibitin1/2, calnexin, and ERas. All four proteins were co-precipitated together with EpCAM-YFP, but not with the control YFP (**Figure 22**). Amongst these four candidates, ERas appeared particularly interesting because of its association to ESCs. Distinctively from Prohibitins 1 and 2, and Calnexin, embryonic stem cell-expressed RAS (ERas), which encodes a Ras-like GTPase protein, was initially characterized as an ESCs-associated molecule that is involved in the tumorigenicity of mouse ESCs (Takahashi, Mitsui and Yamanaka, 2003). Through binding to phosphatidylinositol 3 kinase (PI3K) and phosphorylation of AKT, ERas promotes the growth of ESCs and activates an ERas–AKT–FOXO1 signalling pathway during somatic cell reprogramming (Takahashi, Mitsui and Yamanaka, 2003; Yu *et al.*, 2014). Hence, these data suggested a role for ERas in cell proliferation during early mouse embryonic development (Rodriguez-Viciano *et al.*, 1994; Bedzhov *et al.*, 2014). These aspects qualified ERas as an interesting interaction partner of EpCAM that might explain the cellular effects of EpCAM in ESCs.

#### **4.1.6.2 Function of EpCAM/ERas/Akt signaling in ESCs**

The stable expression of exogenous EpCAM in E14TG2 $\alpha$  ESCs induced an increase in AKT phosphorylation at serine473 and a hyper-activation of AKT under insulin-like growth factor treatment (Sarrach *et al.*, 2018). In opposite, knockout of EpCAM in ESCs reduced the activating phosphorylation of AKT by 72.5% in average, and reduced ERas expression was observed concurrently (Sarrach *et al.*, 2018), suggesting a novel EpCAM/ERas/pAKT

signaling cascade in ESCs.



**Fig. 23: Function of the EpCAM/ERas/Akt signaling in the differentiation of ESCs into cardiomyocytes**

E14TG2 $\alpha$  WT ESCs and stable pCAG (control cell line), Flag-ERas, and myrAKT1 transfectants of E14TG2 $\alpha$  ESCs were cultured under pluripotency conditions in the presence of LIF. (A, B) Cells were harvested and cell lysates (30ug) were used for immunoblotting with ERas- and pAKT-specific antibodies. (C) EBs were generated from E14TG2 $\alpha$  WT ESCs and stable pCAG (control cell line), Flag-ERas, and myrAKT1 transfectants of E14TG2 $\alpha$  ESCs, and were spontaneously differentiated until day 10. The percentage of contraction of EBs were recorded and normalized to 100%. Shown are mean and standard deviations of  $n = 3$  independent experiments. p-value were calculated with a One-way ANOVA test with multiple posthoc testing and Bonferroni correction. \*\*\*\*<0.0001.

To address the function of ERas/Akt signaling in cardiomyocyte formation, a FLAG-tagged

version of ERas and a myristoylated, hyperactive variant of AKT (myrAKT) were stably transfected into E14TG2 $\alpha$  ESCs (**Figure 23A, B**). Ectopic expression of Flag-ERas and myrAKT reduced the contraction rate of EBs by 35% and 40% respectively (**Figure 23C**), suggesting that the over-expression of ERas impairs cardiomyocyte formation *via* ERas/Akt signaling.

Taken all results above, EpCAM expression is tightly regulated during differentiation of ESCs in order to achieve a mandatory spatiotemporal cellular heterogeneity of EpCAM in endo- and mesodermal lineages. Four out of six ESCs EpCAM<sup>-</sup> single clones only partially developed through the mesodermal differentiation and progress was blocked at a Mesp1<sup>+</sup> stage, ultimately leading to a dysfunction of the formation of contracting cardiomyocytes. Potential binding partners and down-stream signaling of EpCAM were further investigated to understand the retained contracting capacity of the remaining two clones. SILAC and Co-IP experiments both identify a binding partner ERas which play a role in ESCs differentiation. The two EpCAM<sup>-</sup> contracting clones display a further impairment of cardiomyocyte formation after ERas knockout, suggesting a complementary function of ERas to cardiomyocyte formation. In addition, loss of EpCAM in ESCs reduce both ERas and phosphorylated AKT expression, suggesting a novel EpCAM/ERas/pAKT signaling pathway.

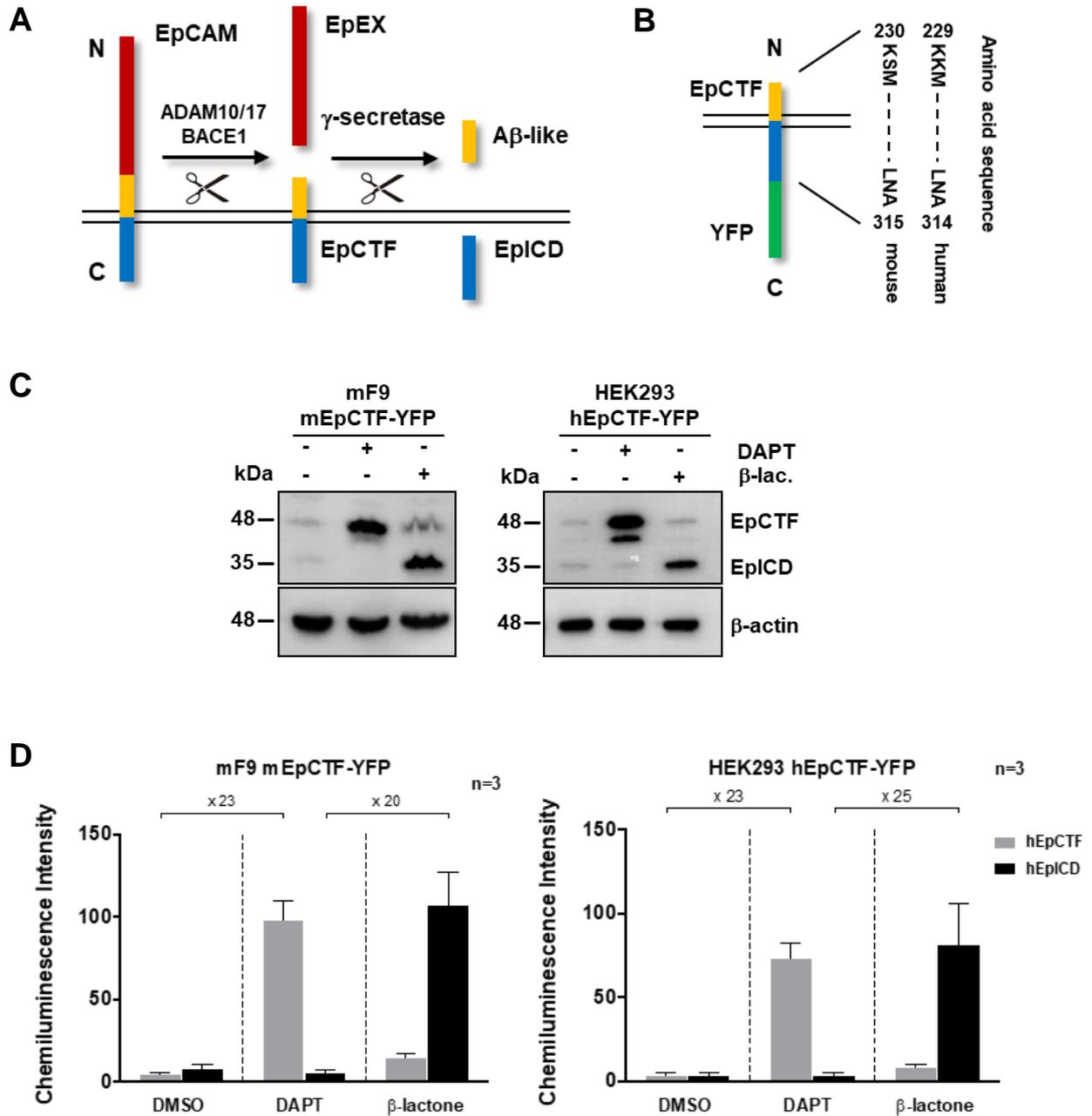
## 4.2 Timing of EpCAM regulation by RIP

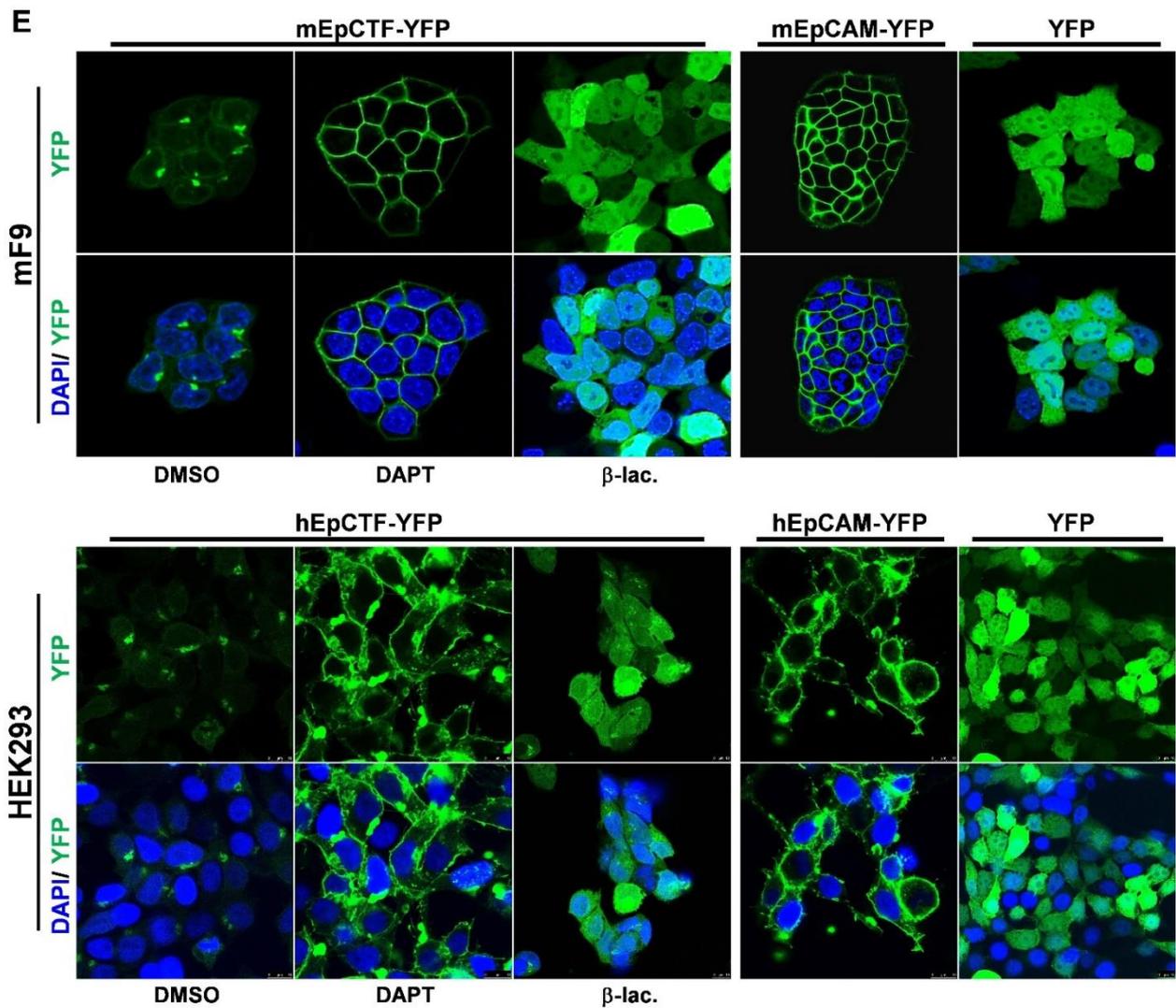
The differential regulation of EpCAM expression during stem cell differentiation was described in section 4.1 of this study. ESCs downregulate the expression of EpCAM starting from day 5.0 of spontaneous differentiation at the mRNA level, while EpCAM protein at the cell membrane was already entirely lost in selected cells at an earlier time point of 3.5-4.0 days (Sarrach *et al.*, 2018). Based on a half-life of EpCAM of 21 h determined in human carcinoma cells, the combination of all data suggested an additional post-translational regulation of EpCAM expression during ESCs differentiation. To the best of our current knowledge, cells have two major post-translational mechanisms to down-regulate EpCAM from the cell membrane, which are cleavage by RIP and endocytosis. Therefore, the timing of EpCAM regulation by RIP was addressed experimentally in this study.

### 4.2.1 Establishment of an EpCTF-YFP cell model to study the pace of intramembrane proteolysis of EpCAM

EpCAM is cleaved by BACE1 or ADAM10/17 in the extracellular domain to release the ectodomain EpEX. The resulting C-terminal fragment (EpCTF) is a substrate that is recognized and further cleaved by the  $\gamma$ -secretase complex. This generates an A $\beta$ -like fragment and an intracellular domain EpICD in succession (**Figure 24A**). With the aim of investigating the efficiency and speed of EpCTF cleavage by  $\gamma$ -secretase, *i.e.* the second cleavage of EpCAM in the process of RIP, the mouse and human variants of the EpCTF that contain the signal peptide, a c-Myc tag, 35 membrane-proximal amino acids (aa) of extracellular domain, and the transmembrane and intracellular domains of EpCAM were fused to the yellow fluorescence protein (YFP), as shown in **Figure 24B**. The resulting EpCTF-YFP molecules (mEpCTF-YFP and hEpCTF-YFP) represent each a mimic of mouse and human EpCTF that allow for the evaluation of the proteolysis of EpCTF by  $\gamma$ -

secretase in living cells through the detection of YFP fluorescence.





**Fig. 24: Generation of mouse and human EpCTF variants**

(A), Scheme of RIP of EpCAM through ADAM10/17, BACE1, and  $\gamma$ -secretase with the releasing fragments (EpEX: EpCAM extracellular domain; EpICD: EpCAM intracellular domain; EpCTF: EpCAM C-terminal fragment). (B), EpCTF-YFP variants include the signal peptide of mouse or human EpCAM (1–23), a short linker peptide, which contains 2 amino acids (KL), the CTF fragment of mouse EpCAM (251–315) and human EpCAM (250–314), as well as the yellow fluorescence protein (YFP). (C), Immunoblotting served to analyze EpCTF-YFP and EpICD-YFP expression with antibody specific to YFP in stable mF9 mEpCTF-YFP transfectants and in stable HEK293 hEpCTF-YFP transfectants. In order to visualize EpCTF-YFP and EpICD-YFP each transfectants was treated as indicated with  $\gamma$ -secretase inhibitor DAPT or proteasome inhibitor  $\beta$ -lactone ( $\beta$ -lac.). The staining of  $\beta$ -actin helped to verify equal protein loading. Shown are representative results of  $n = 3$  independent experiments. (D), Quantification of immunoblotting results in C was obtained from  $n = 3$  independent experiments. Shown are mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate the p-value. (E), mF9 and HEK293 cells with mouse and human EpCTF-YFP, respectively, were treated as indicated with DMSO,  $\beta$ -lactone, or DAPT. YFP fluorescence signal was monitored by laser scanning confocal microscopy. mF9 and HEK293 cells

expressing a YFP-tagged human or mouse full-length EpCAM or YFP served as controls. Shown are the representative results of n = 3 independent experiments.

In order to confirm the correct cleavage and degradation of EpCTF-YFP and EpICD-YFP through  $\gamma$ -secretase and the proteasome, respectively, immunoblotting of whole cell lysates of stable transfectants of mouse EpCTF-YFP in mouse F9 teratoma cells and of human EpCTF-YFP in human embryonic kidney (HEK293) cells was performed. The  $\gamma$ -secretase inhibitor DAPT was applied to treat both cell lines in order to block EpCTF-YFP cleavage, while the proteasome inhibitor  $\beta$ -lactone was applied to block the following degradation of EpICD. Control-treated mF9 and HEK293 transfectants expressed only low levels of mouse and human EpICD-YFP and EpCTF-YFP fragments, respectively (**Figure 24C**). After treatment with DAPT for 24 h, EpCTF-YFP was detected as a minor band of 42kDa together with a dominant band of 45kDa, while EpICD was not or only faintly detected (**Figure 24C**).  $\beta$ -lactone treatment was performed for 12 h to inhibit the proteasome and resulted in the accumulation of both mouse and human EpICD-YFP as 32kDa protein (**Figure 24C**). Three independent immunoblot experiments were implemented to quantify the expression levels of mouse and human EpCTF and EpICD. By using  $\gamma$ -secretase inhibitor DAPT to treat cells to stabilize mEpCTF and hEpCTF, both EpCTFs exhibited a 23-fold increased expression compared with control cells treated with DMSO (**Figure 24D**). By using the proteasome inhibitor  $\beta$ -lactone, mEpICD and hEpICD showed a 20-fold and 25-fold increase compared with DAPT-treated cells, respectively (**Figure 24D**). From these experiments, it can be concluded that mouse and human EpCTF-YFP proteins are processed by  $\gamma$ -secretase, and that the resulting EpICD fragments become degraded *via* the proteasome.

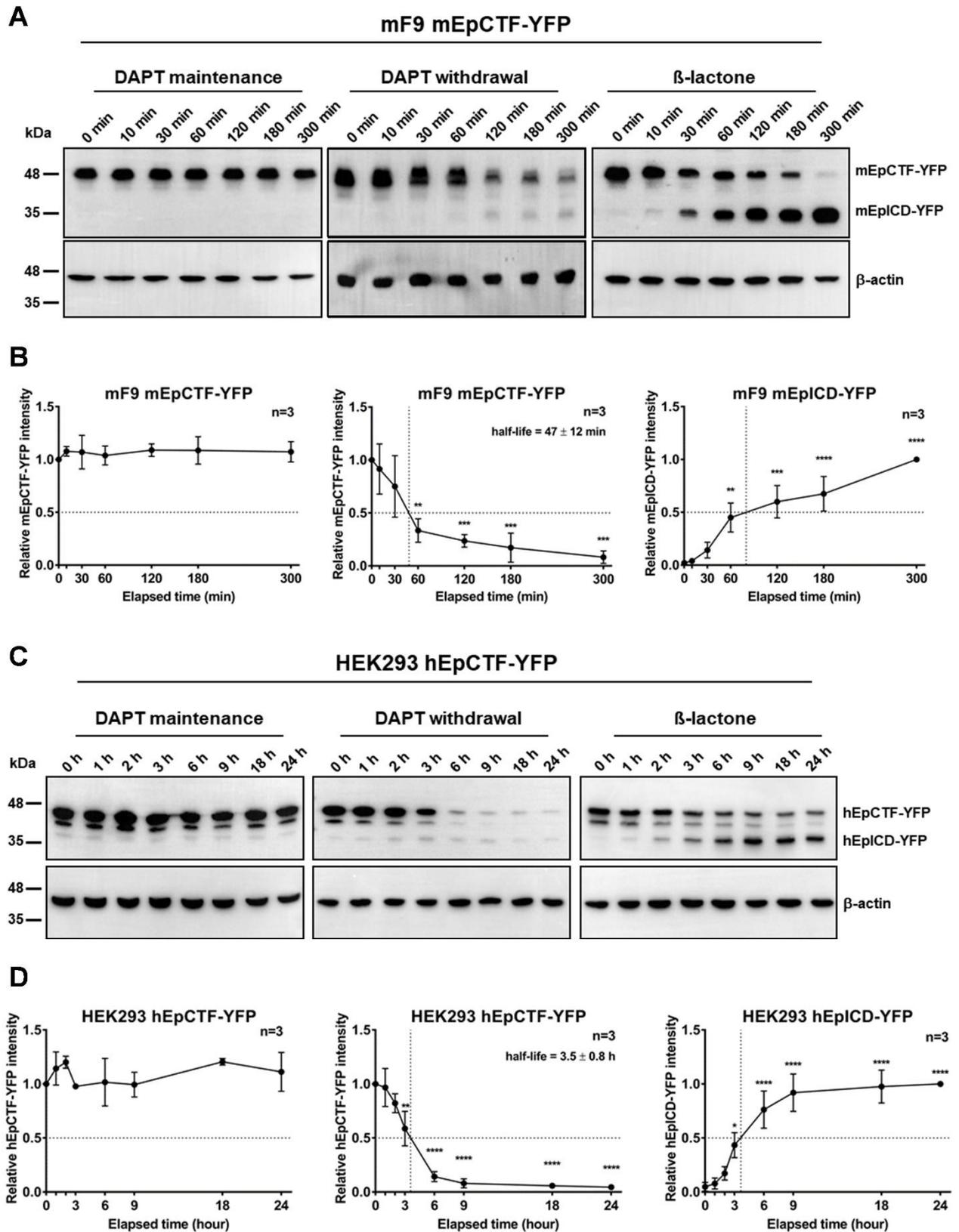
Laser scanning confocal microscopy was then applied to evaluate the correct sub-cellular localization of the EpCTF-YFP fusions. DAPT, DMSO, or  $\beta$ -lactone was used to treat mF9 and HEK293 with mouse and human EpCTF-YFP. After treatment with DMSO, YFP

signals at the plasma membrane could barely be detected, and perinuclear staining and intracellular aggregates were faint (**Figure 24E**). In cells treated with DAPT, mouse and human EpCTF-YFP were readily observed at the plasma membrane. After incubating with  $\beta$ -lactone without DAPT, mouse and human EplCD-YFP exhibited a homogeneous accumulation in the cytoplasm (**Figure 24E**). The control cell lines, which expressed mouse and human full-length EpCAM fused to YFP (EpCAM-YFP), exhibited the expected localization at the plasma membrane. As a further control, YFP-expressing cells displayed an evenly distributed signal throughout the cell (**Figure 24E**).

To sum up, mouse and human EpCTF-YFP variants were expressed in mF9 and HEK293 cells, respectively, and displayed a correct localization and response to the inhibitors compared with endogenous EpCTF (Maetzel *et al.*, 2009).

#### **4.2.2 Biochemical evaluation of the cleavage of EpCTF via $\gamma$ -secretase**

In order to calculate the protein turnover of EpCTF (half-life: 50% turnover) by  $\gamma$ -secretase and to quantify the degradation efficiency of EplCD by the proteasome, time course experiments were carried out in combination with a detection of EpCTF-YFP and proteolytic products by immunoblotting.



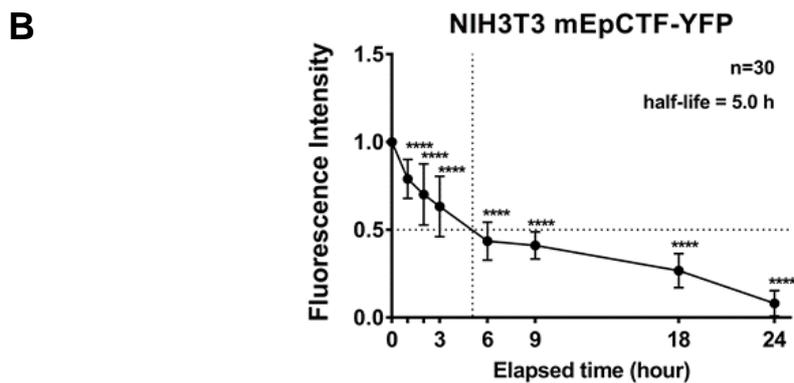
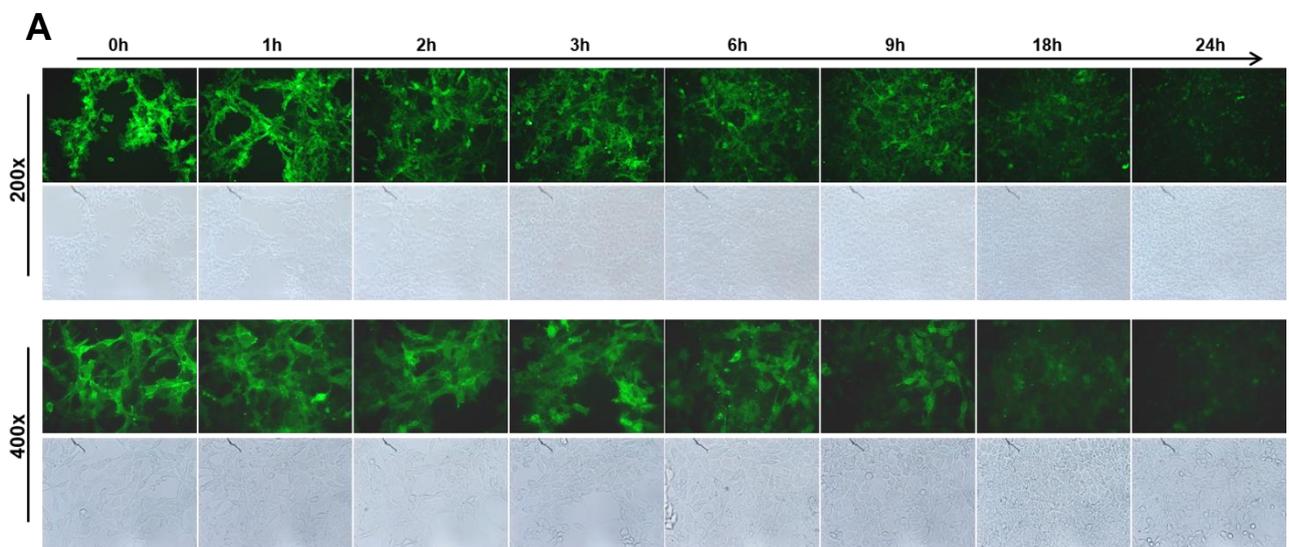
**Fig. 25: Biochemical assessment of mouse and human EpCTF-YFP cleavage**

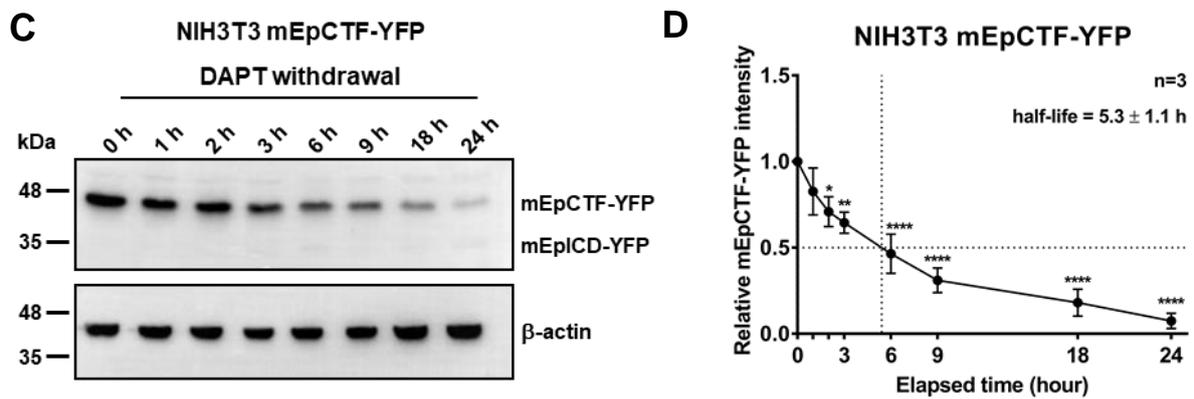
(A), Mouse F9 teratoma cells stably expressing mEpCTF-YFP were treated with DAPT for 12 h. Thereafter, cells were either maintained in medium containing DAPT (DAPT maintenance), or were washed and maintained in normal medium (DAPT withdrawal), or were washed and subsequently maintained in medium containing  $\beta$ -lactone over a total time period of 5 h ( $\beta$ -lactone). Whole cell lysates

of cells at the indicated time points of treatment were analyzed by immunoblotting with GFP-specific antibody, which detects YFP equally well. mEpCTF-YFP and mEpICD-YFP were detected. Shown are representative results of  $n = 3$  independent experiments. (B), Expression of mEpICD-YFP and mEpCTF-YFP was quantified, and mEpCTF 50% protein turnover was calculated based on the  $n = 3$  independent experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\* $<0.01$ ; \*\*\* $<0.001$ ; \*\*\*\* $<0.0001$ . (C), Human HEK293 cells stably expressing hEpCTF-YFP were treated with DAPT for 12 h. Thereafter, cells were either maintained in medium containing DAPT (DAPT maintenance), or were washed and maintained in normal medium (DAPT withdrawal), or were washed and subsequently maintained in medium containing  $\beta$ -lactone over a total time period of 5 h ( $\beta$ -lactone). Whole cell lysates at the indicated time points of treatment were analyzed by immunoblotting with GFP-specific antibody, which detects YFP equally well. hEpCTF-YFP and hEpICD-YFP were detected. Shown are representative results of  $n = 3$  independent experiments. (D), Expression of hEpICD-YFP and hEpCTF-YFP was quantified, and hEpCTF 50% protein turnover was calculated based on  $n = 3$  independent experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate the p-value. \*\* $<0.01$ ; \*\*\* $<0.001$ ; \*\*\*\* $<0.0001$ .

mF9 and HEK293 transfectants expressing mEpCTF-YFP and hEpCTF-YFP, respectively, were treated with DAPT for 12 h in order to inhibit CTF cleavage by  $\gamma$ -secretase. Following this treatment, cells were either maintained in the presence of DAPT in the cell culture medium, or were washed and cultured in standard medium without DAPT, or, as a third variant, were washed and cultured in the presence of the proteasome inhibitor  $\beta$ -lactone. These three different treatments were conducted over a period of 5 h or 24 h for mF9 and HEK293 cells, respectively. Cell samples were collected at different time points. mEpICD-YFP and mEpCTF-YFP expression levels were assessed after 0, 10, 30, 60, 120, 180, and 300 min. In the continuous presence of DAPT, mEpCTF-YFP expression levels was not significantly changed over the observation time period of five h 5 h (**Figure 25A and B**, left panels). Withdrawal of DAPT resulted in mEpCTF-YFP cleavage with a calculated 50% reduction of protein expression at  $47 \pm 12$  min (**Figure 25A and B**, middle panels). After additional treatment with  $\beta$ -lactone, quantification of immunoblots for mEpICD-YFP and mEpCTF-YFP showed that mEpICD-YFP accumulation reached 50% of EpCTF-YFP levels after 70 min (**Figure 25A and B**, right panels).

Cleavage of human EpCTF-YFP and proteasomal degradation of human EpICD-YFP was assessed in stable HEK293 transfectants following the same procedure. No obvious changes could be seen in the expression levels of hEpCTF-YFP after 24 h maintenance in culture medium supplied with DAPT (**Figure 25C and D**, left panels). Following the withdrawal of DAPT, hEpCTF-YFP was cleaved by  $\gamma$ -secretase with a calculated 50% reduction after  $3.5 \pm 0.8$  h (**Figure 25C and D**, middle panel). Inhibition of the proteasome with  $\beta$ -lactone disclosed a 50% stabilization of hEpICD-YFP after 3.5 h in average (**Figure 25C and D**, right panels).





**Fig. 26: Mouse EpCTF-YFP cleavage in mouse NIH3T3 fibroblast cells**

(A), Mouse EpCTF-YFP was stably expressed in NIH3T3 fibroblasts. Stable transfectants were then treated with  $\gamma$ -secretase inhibitor DAPT for 24 h and then cells were cultured in medium without DAPT for another 24 h. Immunofluorescence microscopy was applied to monitor YFP fluorescence at the indicated time points. Shown are the representative pictures from  $n = 3$  independent experiments at 400x and 200x magnification. (B), Quantification of immunofluorescence microscopy results shown in A was carried out from a total of  $n = 30$  cells from  $n = 3$  independent experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\*\*\*<0.0001. (C), Expression of mEpCTF-YFP was visualized by immunoblotting in whole cell lysates of NIH3T3 fibroblasts stably transfected with mEpCTF-YFP using GFP-specific antibody. Similar loading of each sample was confirmed by staining with  $\beta$ -actin-specific antibody. Shown are representative results from  $n = 3$  independent experiments. (D), Quantification of the protein turnover of mEpCTF (50% protein remaining) was calculated based on the  $n = 3$  independent immunoblot experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\*<0.01; \*\*\*<0.001; \*\*\*\*<0.0001.

With a time of  $47 \pm 12$  min and  $3.5 \pm 0.8$  h required to catalyze 50% of the protein amount of mEpCTF-YFP and hEpCTF-YFP by  $\gamma$ -secretase, respectively, this process of enzymatic cleavage is very slow. In order to confirm that the abovementioned values are not specific to the cell type initially used, mEpCTF-YFP and hEpCTF-YFP were expressed in mouse NIH3T3 fibroblasts and in the squamous cell carcinoma line FaDu, respectively.

Based on a 24 h observation period by epifluorescence microscopy, the fluorescence of mEpCTF-YFP was gradually reduced in NIH3T3 cells (**Figure 26A**). The fluorescence intensity was quantified over time in  $n = 10$  cells of  $n = 3$  independent experiments ( $n = 30$  total cells), and the average half-life of mEpCTF-YFP was calculated as 5.0 h (**Figure**

**26B**). The mEpCTF-YFP cleavage over time was confirmed through biochemical analysis of whole cell lysates of mEpCTF-YFP-expressing NIH3T3 fibroblasts at identical time points (**Figure 26C**). mEpCTF-YFP showed a 50% protein turnover rate of  $5.3 \pm 1.1$  h in the NIH3T3 cells, as determined by the immunoblot quantification (**Figure 26D**). Hence, cleavage of mEpCTF-YFP by  $\gamma$ -secretase is a slow process, independently of the cell line used.

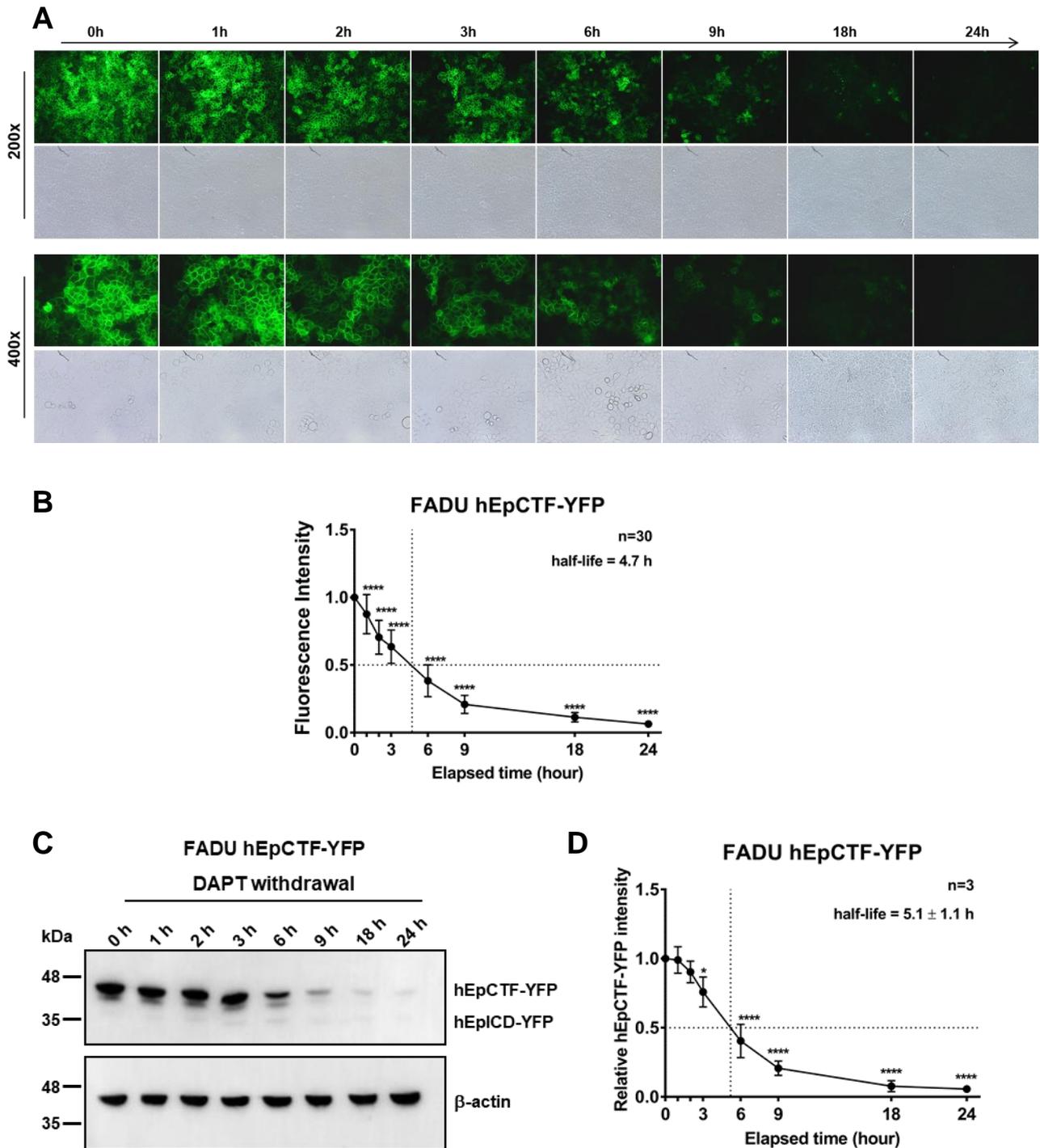


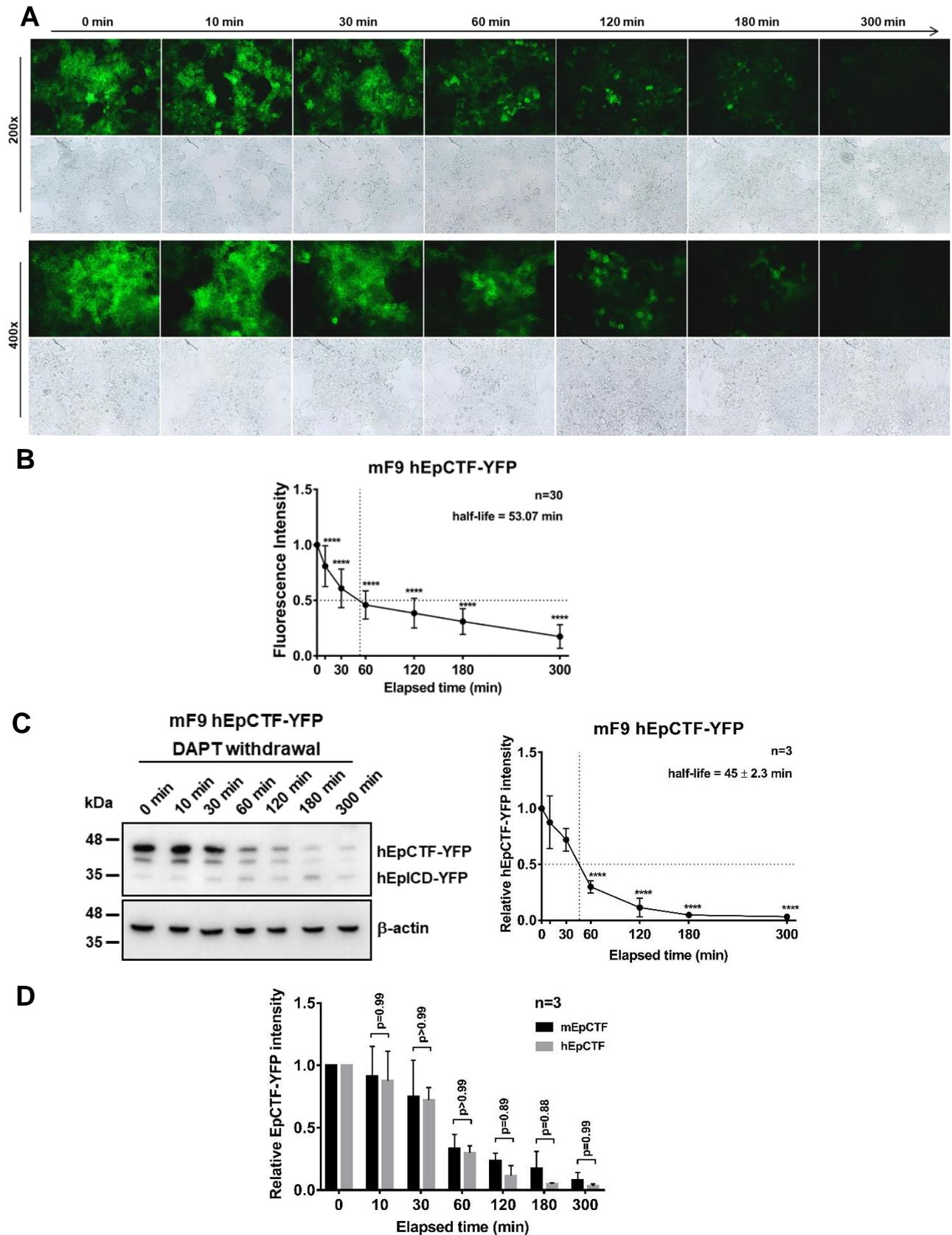
Fig. 27: Human EpCTF cleavage in human FaDu carcinoma cells

(A), Human EpCTF-YFP was stably expressed in FaDu carcinoma cells. Stable transfectants were treated with  $\gamma$ -secretase inhibitor DAPT for 24 h and then cells were maintained in culture medium without DAPT for another 24 h. Immunofluorescence microscopy was applied to monitor YFP fluorescence at the indicated time points. Shown are the representative pictures from  $n = 3$  independent experiments at 400x and 200x magnification. (B), Quantification of immunofluorescence microscopy results shown in A was carried out from a total of  $n = 30$  cells in  $n = 3$  independent experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\*\*\* $<0.0001$ . (C), Expression of hEpCTF-YFP was visualized by immunoblotting in whole cell lysates of FaDu cells stably transfected with hEpCTF-YFP using GFP-specific antibody. Similar loading of each sample was confirmed by staining with  $\beta$ -actin-specific antibody. Shown are representative results from  $n = 3$  independent experiments. (D), Quantification of the protein turnover of hEpCTF (50% protein remaining) was calculated based on  $n = 3$  independent immunoblot experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\* $<0.01$ ; \*\*\* $<0.001$ ; \*\*\*\* $<0.0001$ .

Similar results were obtained with respect to the cleavage of human EpCTF-YFP upon stable transfection in FaDu cells. After assessing the fluorescence of YFP, an average half-life of 4.7 h was calculated (**Figure 27A-B**). According to the biochemical evaluation of hEpCTF-YFP cleavage, hEpCTF-YFP exhibited a protein turnover of 50% at  $5.1 \pm 1.1$  h (**Figure 27C-D**). Hence, cleavage of EpCTF-YFP by  $\gamma$ -secretase is a slow process, independently of the cell line used and irrespective of the species of origin of EpCAM.

#### 4.2.3 The pace of proteolysis of EpCTF is dictated by $\gamma$ -secretase

Despite of the generally very slow proteolysis of mEpCTF and hEpCTF, there was an obvious difference in the 50% protein turnover rates of the EpCTFs in all tested cell lines (minimum: 45 min; maximum: 5.3 h in mF9 cells and HEK293 cells, respectively).



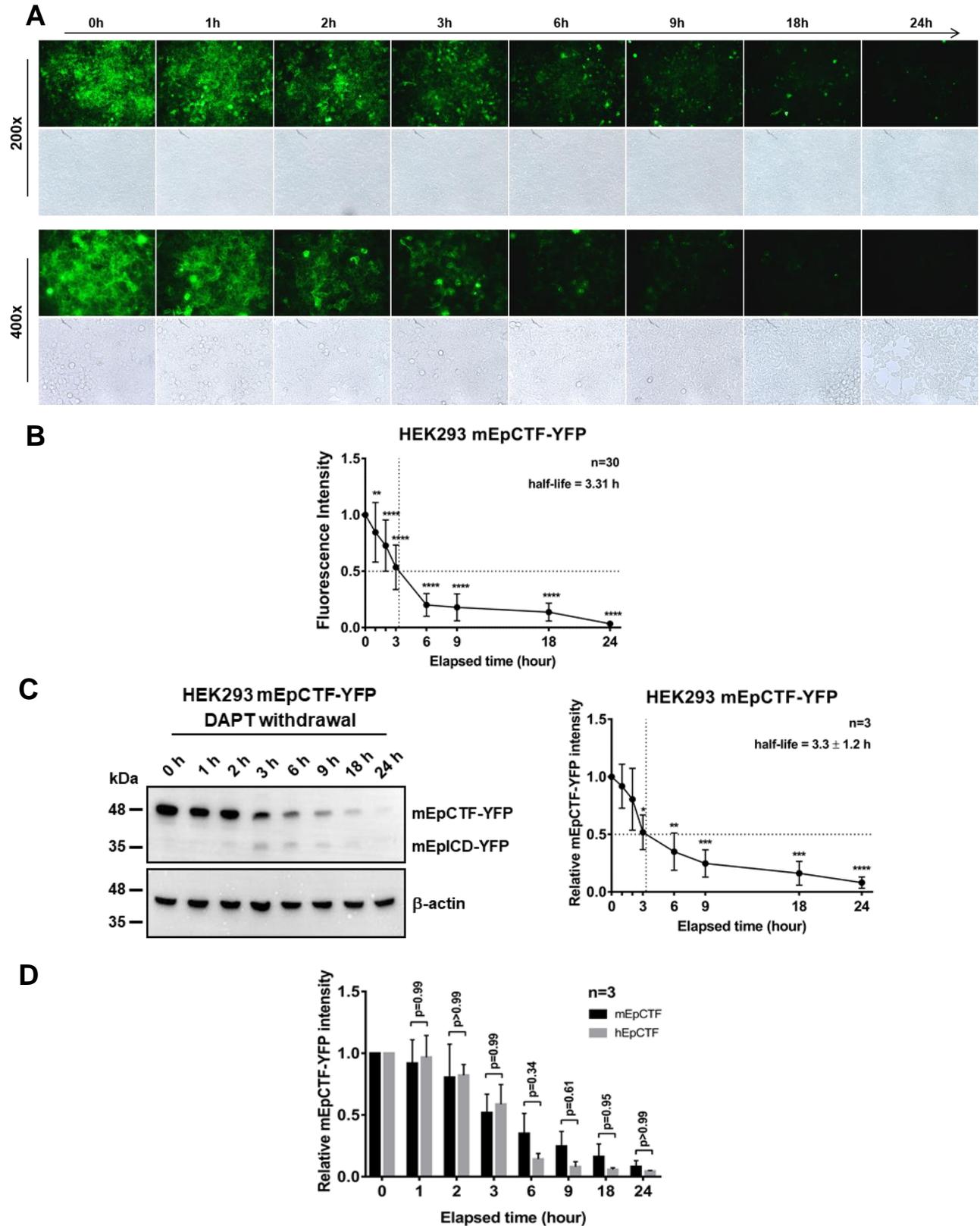
**Fig. 28: hEpCTF cleavage in mouse F9 teratoma cells**

(A), Human EpCTF-YFP was stably expressed in mouse F9 teratoma cells. Stable transfectants were treated with  $\gamma$ -secretase inhibitor DAPT for 24 h and then cells were cultured in medium without DAPT for another 5 h. Immunofluorescence microscopy was applied to monitor YFP fluorescence at the

indicated time points. Shown are the representative pictures from  $n = 3$  independent experiments at 400x and 200x magnification. (B), Quantification of immunofluorescence microscopy results shown in A was carried out from a total of  $n = 30$  cells in  $n = 3$  independent experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\*\*\* $<0.0001$ . (C), Expression of hEpCTF-YFP was visualized by immunoblotting in whole cell lysates of mF9 cells stably transfected with hEpCTF-YFP using GFP-specific antibody. Similar loading of each sample was confirmed by staining with  $\beta$ -actin-specific antibody. Shown are representative results from  $n = 3$  independent experiments. (D), Quantification of the protein turnover of hEpCTF (50% protein remaining) was calculated based on  $n = 3$  independent immunoblot experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\* $<0.01$ ; \*\*\* $<0.001$ ; \*\*\*\* $<0.0001$ . (E), The difference between the half-life of hEpCTF and mEpCTF at each time point of analysis in mF9 cells was calculated with a Two-way ANOVA with multiple *posthoc* testing and Bonferroni correction. The test showed an overall p-value of 0.96 with individual p-value in the range of 0.88-0.99.

In order to define the rate-limiting elements during the EpCTF cleavage, we implemented cross-species swapping experiments, where mEpCTF and hEpCTF-YFP -YFP were expressed in HEK293 and mouse F9 cells, respectively. Epifluorescence microscopy served to record the cleavage of hEpCTF-YFP in mF9 cells, and demonstrated that YFP fluorescence decreased in a time frame of 30 to 60 min following DAPT withdrawal (**Figure 28A**). Quantification of the fluorescence of YFP over time in  $n = 10$  cells of the  $n = 3$  independent experiments ( $n = 30$  total cells), the average half-life period was calculated as 53.07 min (**Figure 28B**). Based on the immunoblot analysis of hEpCTF-YFP in mouse F9 cells, the cleavage rate detected *via* fluorescence imaging could be confirmed. The protein turnover reached 50% at  $45 \pm 2.3$  min (**Figure 28C**). On that account, according to microscopy as well as immunoblotting experiments, the cleavage rate of human and mouse EpCTF-YFP was very similar in mF9 cells, and 50% of hEpCTF-YFP molecules underwent cleavage at approx. 45 min in mF9 cells as compared to 3.5 h in HEK293 cells. These similar rates of cleavage were corroborated by statistical analysis with a Two-way ANOVA of each time point between hEpCTF and mEpCTF in mF9 cells. With an overall p-

value of 0.96, the cleavage rate of human and mouse EpCTF did not differ (**Figure 28D**).



**Fig. 29: mEpCTF cleavage in human HEK293 cells**

(A), Mouse EpCTF-YFP was stably expressed in human HEK293 cells. Stable transfectants were then treated with  $\gamma$ -secretase inhibitor DAPT for 24 h and then cells were cultured in medium without DAPT for another 24 h. Immunofluorescence microscopy was applied to monitor YFP fluorescence at the

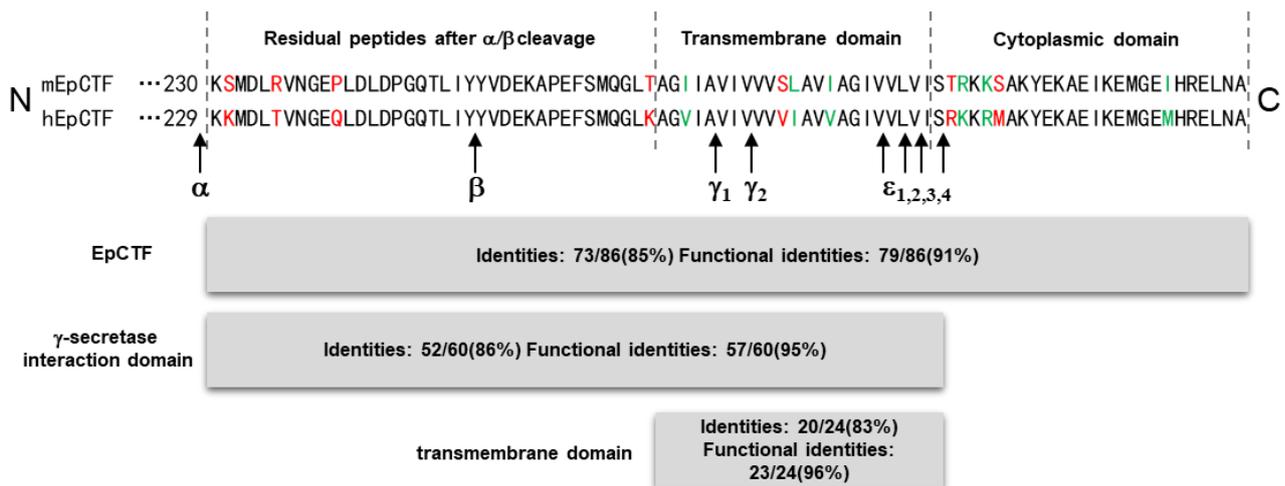
indicated time points. Shown are the representative pictures from  $n = 3$  independent experiments at 200x and 400x magnification. (B), Quantification of immunofluorescence microscopy results shown in A was carried out from  $n = 30$  cells in  $n = 3$  independent experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\*\*\* $<0.0001$ . (C), Expression of mEpCTF-YFP was visualized by immunoblotting in whole cell lysates of mF9 cells stably transfected with mEpCTF-YFP using GFP-specific antibody. Similar loading of each sample was confirmed by staining with  $\beta$ -actin-specific antibody. Shown are representative results from  $n = 3$  independent experiments. (D), Quantification of the protein turnover of mEpCTF (50% protein remaining) was calculated based on  $n = 3$  independent immunoblot experiments. Shown are the mean values with standard deviations. One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\* $<0.01$ ; \*\*\* $<0.001$ ; \*\*\*\* $<0.0001$ . (D), The difference between the half-life of hEpCTF and mEpCTF at each time point of analysis in mF9 cells was calculated with a Two-way ANOVA with multiple *posthoc* testing and Bonferroni correction. The test showed an overall p-value of 0.10 with individual p-value in the range of 0.34-0.99.

Similarly, epifluorescence microscopy in combination with immunoblotting was used to analyze the cleavage pace of mEpCTF-YFP in HEK293 cells. Both assays confirmed that mEpCTF-YFP protein expression was reduced over 24 h and the calculated average 50% protein turnover was 3.31 h, according to the fluorescence intensity kinetics of a total of  $n = 30$  cells (**Figure 29A-B**). Evaluation of the biochemical assessment of the half-life of mEpCTF-YFP in HEK293 cells confirmed a 50% reduction after  $3.3 \pm 1.2$  h (**Figure 29C**). No statistical difference was detected between mEpCTF and hEpCTF in terms of their cleavage pace in HEK293 cells. Two-way ANOVA displayed an overall p-value at 0.10 and individual p-value changed in the range of 0.34 and 0.99 at the different time points of analysis (**Figure 29D**). Therefore, cross-species swapping experiments illustrated the association between the cleavage pace of EpCTF variants and endogenous  $\gamma$ -secretase activity, rather than the species of origin of the substrate EpCTF.

#### **4.2.4 Molecular basis for the observed similarity in cleavage pace of EpCTFs from different species**

The cleavage rate of mouse and human EpCTF variants was determined by  $\gamma$ -secretase, rather than EpCTF itself. Therefore, amino acid (aa) sequences of mEpCTF and hEpCTF

were analyzed regarding their identity, functional similarity, as well as differences in a protein stretch from aa 230-315 of mEpCTF and 229-314 of hEpCTF. mEpCTF and hEpCTF showed a functional identity of 91% in the abovementioned aa stretch, and functional identity reached 95% in reported  $\gamma$ -secretase cleavage sites and 96% in transmembrane domains (Hachmeister *et al.*, 2013; Tsaktanis *et al.*, 2015) (**Figure 30**). EpCTF exhibited identical  $\gamma$ -secretase cleavage sites between mEpCTF and hEpCTF, with exception of the  $\epsilon$ 4 site, which differed from a threonine to an arginine, respectively (**Figure 30**). Hence, EpCTF variants share a high degree of conservation across mouse and human species, particularly in  $\gamma$ -secretase cleavage sites. Therefore, we conclude that the cleavage pace is dependent on the  $\gamma$ -secretase complex rather than the CTF substrate.



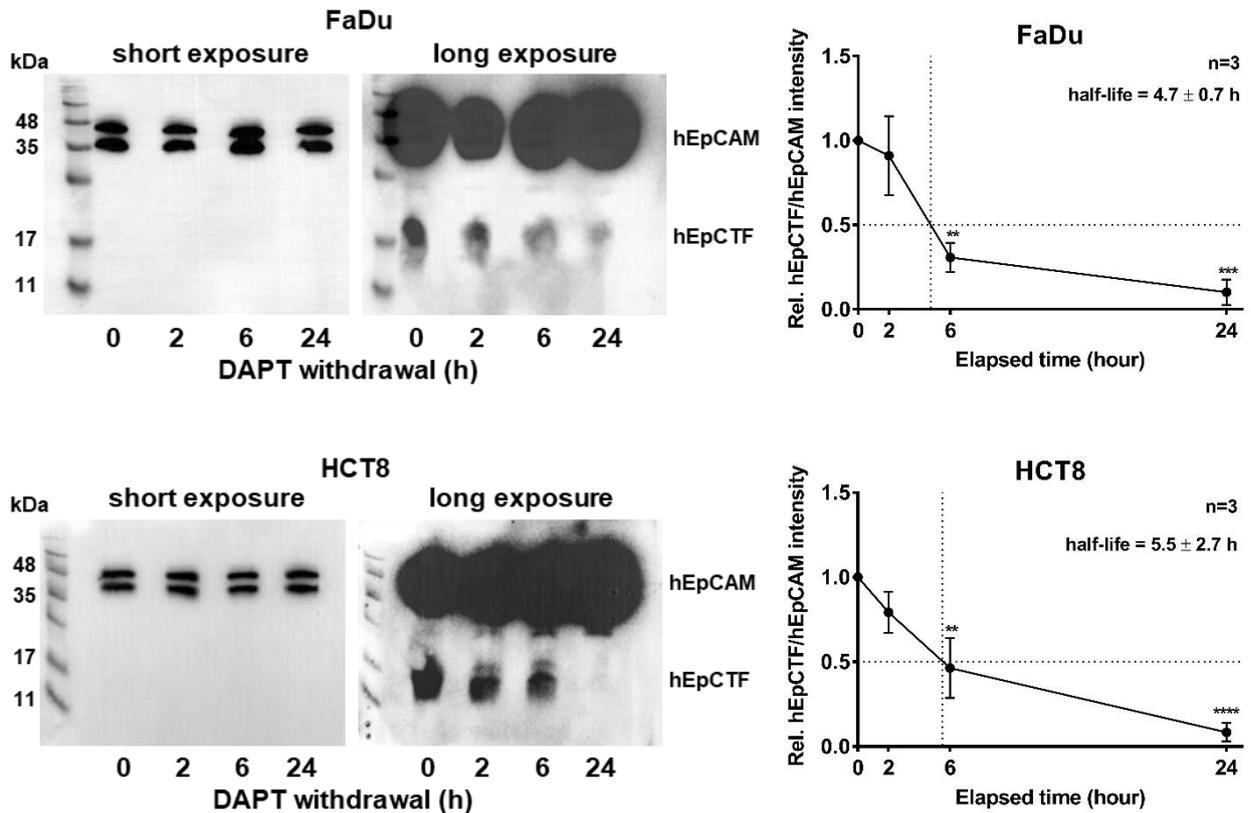
**Fig. 30: Similarities in amino acid sequences of EpCTF variants**

Comparison of the aligned amino acid (aa) sequences of mouse and human EpCTF. AA in red font mark differences between mouse and human EpCTF variants. AA in green font mark residues with different aa, but with functional equivalence. mEpCTF and hEpCTF showed a functional identity of 91%, demonstrating that their amino acid sequences were highly conserved. The  $\gamma$ -secretase cleavage sites and transmembrane domains showed a functional identity of 95% and 96%, respectively. Most  $\gamma$  and  $\epsilon$  cleavage site of the EpCTF variants were identical with the exception of the  $\epsilon$ 4 site.

#### 4.2.5 Assessment of endogenous EpCTF cleavage pace

In the next experiments, we aimed to validate the findings on the slow cleavage rate of

EpCTF-YFP variants for endogenous EpCTF cleavage in membrane fractions of carcinoma cells. DAPT was used to treat the head and neck squamous carcinoma cell line FaDu and the ileocecal adenocarcinoma cell line HCT8 for 24 h, in order to block cleavage of endogenous EpCTF by  $\gamma$ -secretase. Following withdrawal of DAPT, the kinetics of EpCTF cleavage was assessed in immunoblot experiments.



**Fig. 31: Assessment of endogenous EpCTF cleavage**

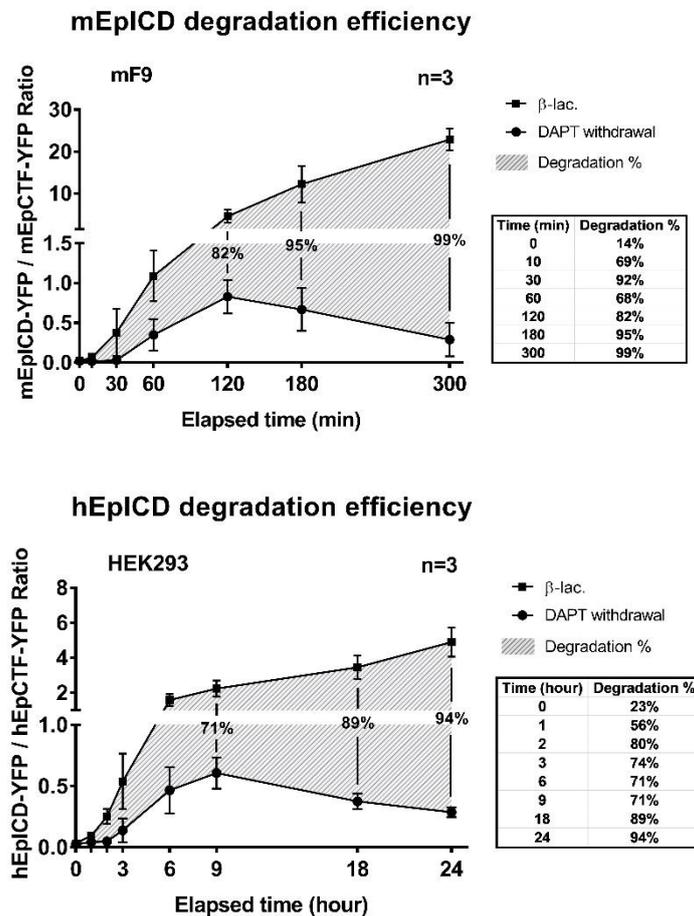
Assessment of endogenous EpCTF cleavage in FaDu and HCT8 cells. FaDu and HCT8 cells were maintained with the  $\gamma$ -secretase inhibitor DAPT for 24 h and then cells were cultured in medium without DAPT for another 24 h. Subsequently, the expression of endogenous EpCTF was monitored with EpICD-specific antibody. Shown are the representative results of  $n = 3$  independent immunoblot experiments (left panels), as well as the mean values of EpCTF expression intensities with standard deviations (right panels). One-way ANOVA with multiple *posthoc* testing and Bonferroni correction served to calculate p-value. \*\*<0.01; \*\*\*<0.001; \*\*\*\*<0.0001.

DAPT treatment of HCT8 and FaDu cells resulted in the accumulation of endogenous human EpCTF fragment, which was subject to a slow cleavage after withdrawal of DAPT (**Figure 31**, left panels). Quantification of immunoblot results disclosed a half-life of endogenous EpCTF of  $4.7 \pm 0.7$  h in FaDu cells and of  $5.5 \pm 2.7$  h in HCT8 cells (**Figure**

31, right panels). Therefore, it was confirmed that the cleavage process of endogenous human EpCTF via  $\gamma$ -secretase was slow in the carcinoma cells, too.

#### 4.2.6 EpICD degradation by the proteasome shows high efficiency

Following cleavage of EpCTF by  $\gamma$ -secretase, EpICD is released into the cytoplasm and can translocate into the nucleus (Maetzel et al., 2009; Chaves-Pérez et al., 2013). Generally, EpICD exhibited a low stability and only a small fraction of it was detected in nucleus (Maetzel et al., 2009; Pan et al., 2018), where EpICD deploys its functions as a signaling active moiety (Maetzel et al., 2009; Lu et al., 2010; H. P. Huang et al., 2011; Denzel et al., 2012; Lin et al., 2012; Yu, Ma and Wang, 2017). In order to control functions of EpICD, cells require modalities of degradation, which are commonly assumed by the proteasome. With the aim to quantify the proteasomal degradation efficiency of human and mouse EpICD, the ratios of EpICD to EpCTF were calculated from western blots of cells treated with and without  $\beta$ -lactone.



### **Fig. 32: Degradation efficiency of EpICD**

Immunoblotting results from 4.2.2 served to calculate the ratios of EpICD-YFP and EpCTF-YFP in order to quantify the degradation efficiency of EpICD. Differences between EpICD/EpCTF ratios in the absence and presence of the proteasome inhibitor  $\beta$ -lactone represent the percentage of degradation of mEpICD and hEpICD variants in the top and bottom panel, respectively. Shown are mean values with standard deviations of each  $n = 3$  independent experiments. Calculated percentages of degradation over time are given in the in-lay table.

In mF9 cells treated with  $\beta$ -lactone in the absence of DAPT, the EpICD/EpCTF ratio exhibited a gradual increase from 0.02 to 22.88, demonstrating the accumulation of mEpICD. In the absence of  $\beta$ -lactone and following the withdrawal of DAPT, the EpICD/EpCTF ratio presented an initial increase to 0.83 and a subsequent decrease to 0.29, suggesting a sequential accumulation of mEpICD followed by a degradation by the proteasome (**Figure 32**, left panel). The differences in ratios between two groups of treatment represented the percentages of mEpICD degradation by the proteasome. Five hours after withdrawal of DAPT, the degradation of mEpICD reached 99% in the absence of  $\beta$ -lactone (**Figure 32**, left panel). Similarly, the degradation of hEpICD reached 94% after 24 h of withdrawal of DAPT in HEK293 cells (**Figure 32**, right panel). Thus, the proteasome can degrade mouse and human EpICD fragments in an efficient manner, which occurs after cleavage of EpCTF by  $\gamma$ -secretase.

## 5. DISCUSSION

The differentiation of pluripotent ESCs into mature cells of the adult organism, which assume a huge diversity of functions, is a highly complex process that is orchestrated by numerous regulatory molecules. These molecules are comprised of cell surface receptors, intracellular mediators, transcription factors, and effector molecules. One cell surface protein that is highly expressed in pluripotent ESCs is the epithelial cell adhesion molecule EpCAM. Throughout ESCs differentiation, EpCAM is characterized by a spatiotemporal regulation, which eventually results in a programmed differential expression in mature cells. Accordingly, in adult organisms EpCAM is primarily found in epithelial cells, and lacking in all other cells.

In the present study, this spatiotemporal regulation of EpCAM in ESCs was analyzed in-depth and functional consequences were addressed.

### 5.1 EpCAM is required for full differentiation of pluripotent ESCs

Pluripotent ESCs express high levels of EpCAM, whereas cellular levels of EpCAM differ substantially during early differentiation of ESCs (González *et al.*, 2009; Sarrach *et al.*, 2018). Here, clusters of EpCAM<sup>+</sup> and EpCAM<sup>-</sup> differentiating ESCs were interspersed after 3.5-4 days of spontaneous differentiation of E14TG2 $\alpha$  ESCs in EBs. Consecutive single staining and double-staining demonstrated that EpCAM was co-expressed with Foxa2. Foxa2 is a transcription factor typically expressed in visceral endodermal and endodermal cell clusters in embryos at a later stage of differentiation (Burtscher and Lickert, 2009b). Similarly, the formation of visceral endoderm at the outer rim of EBs has been confirmed during the early differentiation of ESCs (Liu *et al.*, 2009). In the present study, co-localization of EpCAM with Foxa2 was frequently observed in marginal cells of EBs, suggesting an expression of EpCAM in nascent visceral endoderm.

Vimentin, an intermediate filament protein expressed in mesodermal cells, displayed a

mutually exclusive expression pattern with EpCAM in EBs. Generally, vimentin<sup>+</sup> cells were entirely devoid of EpCAM, which supported the notion that differentiation of mesodermal cells required a strict loss of EpCAM (Sarrach *et al.*, 2018). This assumption was further corroborated at the functional level *via* a forced retention of EpCAM in ESCs through exogenous expression. Overexpression of EpCAM from the strong cytomegalovirus promoter negatively impacted on mesodermal differentiation into cardiomyocytes (Sarrach *et al.*, 2018). However, a genetic knockout of the *EPCAM* gene using the CRISPR/Cas9 technology in E14TG2 $\alpha$  ESCs resulted in inhibitory effects on cardiomyocytes differentiation too, *i.e.* a substantial reduction or loss of EBs contraction upon spontaneous differentiation of EpCAM knockout ESCs. Generally, guided differentiation of EpCAM knockout single cell clones further demonstrated a reduced pluripotency of these cells, which is in accordance with reported functions of EpCAM in human and mouse ESCs, as well as in porcine induced pluripotent stem cells (iPS) (González *et al.*, 2009; Lu *et al.*, 2010; Ng *et al.*, 2010; Yu, Ma and Wang, 2017). Based on EpCAM knockout and overexpressing ESCs clones, it appeared that the proper regulation of EpCAM during differentiation is required to achieve the co-existence of EpCAM<sup>+</sup> and EpCAM<sup>-</sup> cells. As a consequence, both, the entire loss and the forced overexpression resulted in disrupted differentiation.

## **5.2 Mandatory segregation of EpCAM<sup>+</sup> and EpCAM<sup>-</sup> clusters during EBs differentiation**

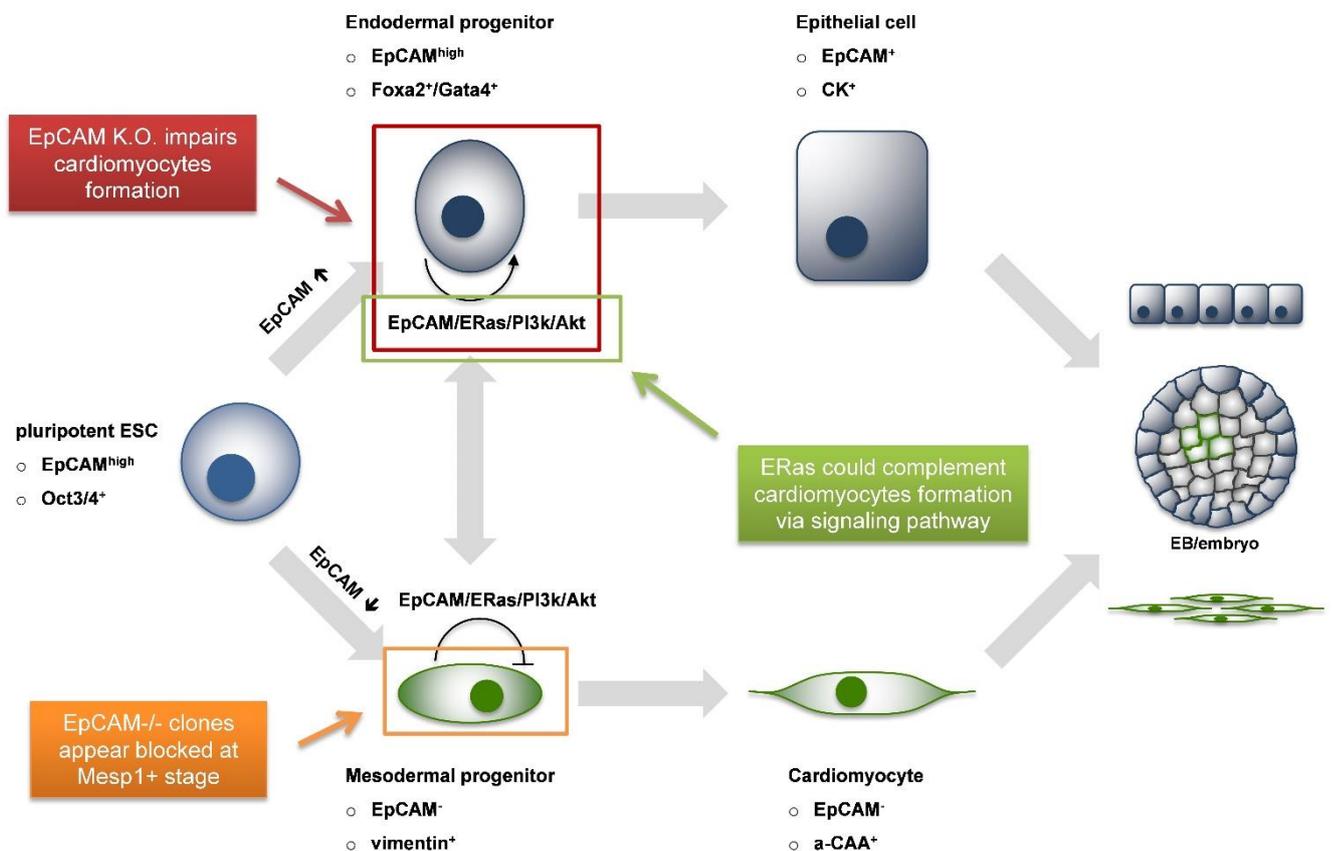
The presented study analyzed the patterns of EpCAM repression during the differentiation of mouse ESCs in a 3D-differentiation model. This EB-based model simulates various aspects of mouse embryogenesis in early stages and combines it with ability to genetically manipulate ESCs (Desbaillets *et al.*, 2000). As such EBs derived from ESCs represent a useful *in vitro* simulation of an embryo comprised of internal meso- and ectoderm lineages, as well as an external visceral/primitive endoderm (Doetschman *et al.*, 1985; Nishikawa,

Jakt and Era, 2007), which involves the forming of a primitive streak together with an anteroposterior axis through Wnt signaling (Nusse *et al.*, 2008; ten Berge *et al.*, 2008). In this model, EpCAM<sup>+</sup> and EpCAM<sup>-</sup> cell clusters were segregated early in spontaneous differentiation, but remained in close proximity. The spatial separation of EpCAM<sup>+</sup> and EpCAM<sup>-</sup> clusters could possibly be due to a direct effect of EpCAM on cell adhesion (Litvinov, Bakker, *et al.*, 1994) or a negative effect on cell-cell contacts mediated by E-cadherin (Litvinov *et al.*, 1997). E-cadherin is a major adhesion molecule of mouse intestinal epithelial cells (Solanas *et al.*, 2011) and during zebrafish gastrulation (Ulrich *et al.*, 2005). In the mouse intestine, E-cadherin-mediated cell-cell adhesion must be abrogated in order to allow the migration of differentiating cells along the developing crypts (Solanas *et al.*, 2011). Therefore, a similar requirement for a loss of EpCAM during the segregation of differentiating cells in EBs is conceivable, based on EpCAM's own adhesion functions and on its effects on E-cadherin.

### **5.3 Loss of EpCAM expression in pluripotent ESCs inhibits cardiomyocytes formation**

Based on the data in the present thesis, EpCAM over-expression imposes an inhibitory effect on the formation of cardiomyocytes, while a CRISPR/Cas9-mediated mutation of EpCAM demonstrated the necessity of endodermal EpCAM<sup>+</sup> cells for the mesodermal lineage to generate cardiomyocytes (**Figure 33**). The expression of EpCAM supported the expression of *Foxa2*, *Gata4*, as well as alpha-fetoprotein (Afp) in endodermal cells derived from ESCs (Sarrach *et al.*, 2018). In line with these findings, it has been reported that the development of cardiomyocytes required a physical contact of mesodermal progenitors with endodermal cells (Foley *et al.*, 2006). More precisely, cardiomyocyte precursors need to be in physical contact with *Gata4*-producing Sox17<sup>+</sup>/EpCAM<sup>+</sup> visceral endoderm, for the proper instruction of cardiomyocytes differentiation (Pal and Khanna, 2005; Holtzinger, Rosenfeld and Evans, 2010). A reported progression of Sox17<sup>+</sup>/EpCAM<sup>+</sup> visceral

endodermal cells to hepatocytic progenitors (Holtzinger, Rosenfeld and Evans, 2010) further illustrated that EpCAM<sup>+</sup> cells were mutually dependent on EpCAM<sup>-</sup> cells for their own differentiation. On that account, a high expression of EpCAM constitutes a major characteristic of hepatocytic stem cells in human (Schmelzer, 2007; Yoon *et al.*, 2011). Furthermore, EpCAM acts as a de-repressor of the Wnt signaling cascade in zebrafish, which licenses cells to mature to functional hepatocytes (Lu *et al.*, 2013). In line with this report, EpCAM over-expression in ESCs facilitated the expression of hepatocytic markers *Afp* and *Fn1* (Sarrach *et al.*, 2018).



**Fig. 33: Schematic illustration of EpCAM expression in differentiating ESCs**

Scheme shows that ESCs differentiated into either EpCAM<sup>+</sup> endodermal progenitors or EpCAM<sup>-</sup> mesodermal progenitors during EBs differentiation. EpCAM<sup>+</sup> endodermal progenitors and EpCAM<sup>-</sup> mesodermal progenitors in EBs were further differentiated into epithelial cells and cardiomyocytes respectively. Repression of EpCAM impairs the cardiomyocytes formation through the physical contact of endodermal and mesodermal progenitors. ERas as a downstream molecule of EpCAM has the capacity of complementing for the loss of EpCAM to support the formation of Gata4<sup>+</sup> endodermal cells. Differentiation of EpCAM<sup>-</sup> ESCs was blocked at Mesp1<sup>+</sup> stage.

#### **5.4 EpCAM knockout ESCs clones are blocked in differentiation at the Mesp1<sup>+</sup> stage**

The transcriptional factor Mesoderm Posterior BHLH Transcription Factor 1 (Mesp1) plays a central role in the development of the human cardiovascular system (Mazzotta *et al.*, 2016b). Accordingly, Mesp1 is required during early stages of the development of cardiomyocytes *via* the regulation of cardiac mesoderm at E6.5. As a result, the first heart field is formed, followed by the heart tube (Spater *et al.*, 2014; Chen *et al.*, 2015). During the migration process to form the heart crescent, down-regulation of Mesp1 by cardiomyocytes progenitors further facilitates the maturation to cardiomyocytes (Paige *et al.*, 2015).

In the present study, Mesp1 expression was greatly increased and Wnt5a was mildly enhanced at day 10 of spontaneous differentiation of EpCAM knockout clones. However, the expression of the central downstream regulators of cardiomyocytes differentiation Gata4, Nkx2.5, and Wnt11, as well as the cardiomyocytes marker  $\alpha$ -CAA was strongly reduced in EpCAM knockout clones at day 10 of spontaneous differentiation. At this time point, down-regulation of Mesp1 is supposed to be accomplished and is a prerequisite for cardiomyocytes differentiation in EBs. Hence, EpCAM knockout clones demonstrated a differentiation block at a Mesp1<sup>+</sup> stage with a subsequent lack of further differentiation into cardiomyocytes (**Figure 33**). Due to a reduction of Wnt11, meso-endoderm cohesion is potentially affected (Ulrich *et al.*, 2005), and mesodermal differentiation is ultimately blocked.

#### **5.5 Participation of the EpCAM/ERas/AKT axis in endo/ mesodermal differentiation**

The observed dual capacity of EpCAM to positively and negatively affect ESCs differentiation, might result from its interaction with the hyperactive Ras GTPase ERas, which reportedly affects the proliferation and teratogenic capacity of ESCs (Takahashi, Mitsui and Yamanaka, 2003; Lee *et al.*, 2009a). Forced expression of EpCAM facilitated

the activating phosphorylation of the serine/threonine kinase AKT at serine<sup>473</sup> in ESCs. In turns, ERas over-expression or activated AKT simulated the inhibitory function of EpCAM during the formation of cardiomyocytes, although with a reduced inhibitory ability compared with EpCAM. Our findings were corroborated by published results demonstrating that a loss of ERas is required in E7.5 embryos to facilitate the formation of the PS, as well as the generation of mesoderm (Zhao *et al.*, 2015). Additionally, they reported on a retention of ERas in the endoderm of the same stage of gestation (Zhao *et al.*, 2015). This finding further supports a role for the EpCAM/ERas signaling axis in the positive regulation of endodermal cells.

EpCAM knockout clones, which showed a retained ability of cardiomyogenesis, revealed to be severely impaired in the formation of cardiomyocytes following an additional knockout of ERas. These findings suggested an ability of ERas to complement for the loss of EpCAM to support the formation of Gata4<sup>+</sup> endodermal cells, which are required by cardiomyocytes progenitors for proper cardiomyocytes differentiation (**Figure 33**). An ERas interactor Galectin-1 has been shown to mediate anchorage of Ras proteins to the cell membrane (Paz *et al.*, 2001) and to induce the mouse ESCs proliferation through the Src/ERas/Akt/mTOR signaling pathway (Lee *et al.*, 2009b). Thus, the interaction of EpCAM with ERas was also likely to recruit the downstream molecules to ERas and induce the ERas/Akt signalling pathway. In line with the notion of an EpCAM/ERas signaling axis that is active in stem cells, EpCAM and ERas both positively impact on the reprogramming efficiency of somatic cells into iPS (Huang *et al.*, 2011; Yu *et al.*, 2014; Kuan *et al.*, 2017). Therefore, EpCAM/ERas/AKT represents a novel signaling cascade in ESCs that supports endodermal cell differentiation, which are in turns required to instruct mesodermal differentiation to cardiomyocytes.

## 5.6 Dynamic EpCAM expression in ESCs differentiation and EMT process

In non-pathological processes such as embryonic development, the formation of mesoderm from the epiblast (primitive ectoderm) occurs through the process of EMT, in which epiblast cells with an epithelial phenotype selectively lose surface proteins mostly involved in cell adhesion, and differentiate into mesenchyme (Rivera-Pérez, Mager and Magnuson, 2003; Mikawa *et al.*, 2004; Migeotte *et al.*, 2010; Nowotschin and A.-K. Hadjantonakis, 2010). During the differentiation of EBs, which simulates early embryonic development *in vitro*, EpCAM expression was retained in endodermal derived cells, but was repressed in mesodermal lineages (Sarrach *et al.*, 2018), suggesting that the regulation of the EpCAM expression in non-pathological process e.g. ESCs differentiation was very likely in the context of EMT. In addition, the EpCAM dynamic expression is important for full differentiation of ESCs as balanced integration of EpCAM<sup>+</sup> and EpCAM<sup>-</sup> cells is required (Sarrach *et al.*, 2018). Accordingly, it was reported that EMT transcription factor Zeb1 can down-regulate the expression of both, E-cadherin and EpCAM, in zebrafish development (Vannier *et al.*, 2013).

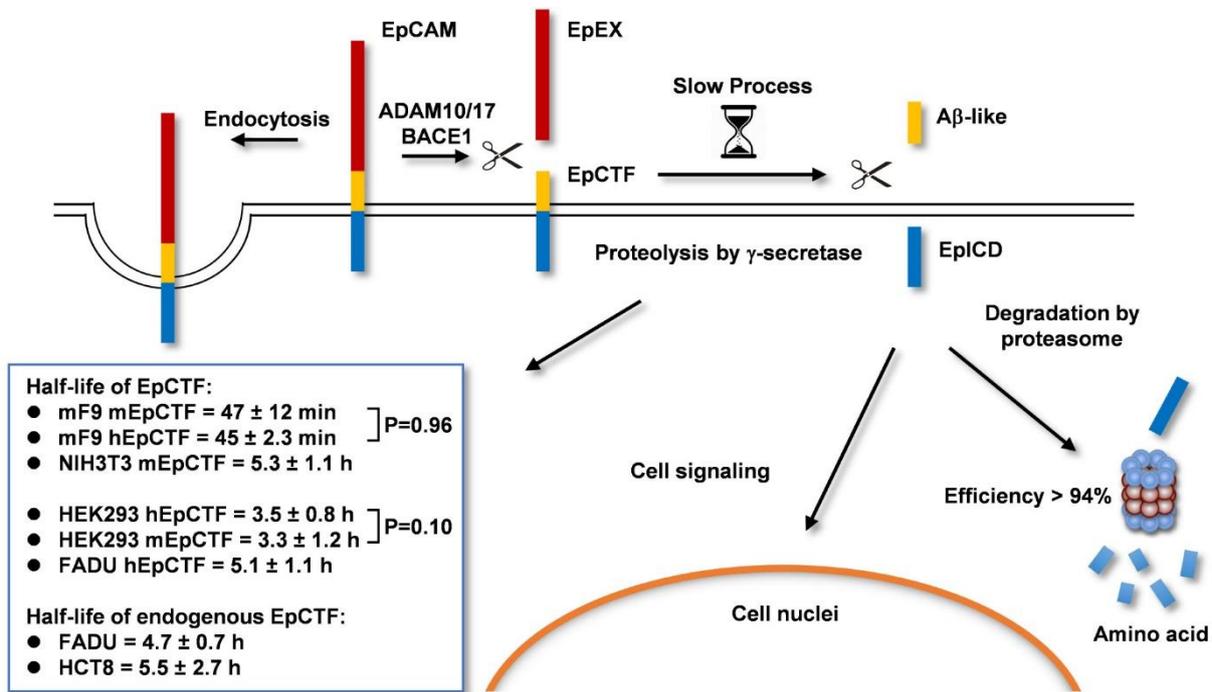
In pathological processes, EpCAM exhibits a dynamic expression in cancer progression and shows a frequent loss in the course of EMT (Gorges *et al.*, 2012; Gires and Stoecklein, 2014; Wang *et al.*, 2016). The EpCAM expression has reportedly served as a valuable marker for the evaluation of EMT heterogeneity, which is strongly associated with cancer progression (Liu *et al.*, 2019). In head and neck cancers (HNSCC), EpCAM was a reliable marker to determine the epithelial differentiation status of cancer cells, as opposed to cells that have undergone a partial EMT, as has been demonstrated through single cell RNA-sequencing of primary tumors and metastases (Puram *et al.*, 2017). Results from our lab confirmed that a loss of EpCAM in HNSCC was correlated with worsened clinical outcome (Baumeister *et al.*, 2018). Moreover, high expression of epidermal growth factor receptor EGFR in combination with low levels of EpCAM were associated with very poor prognosis

in HNSCCs, whereas EGFR<sup>low</sup>/EpCAM<sup>high</sup> HNSCCs were defined by an outstandingly good prognosis (Pan *et al.*, 2018). An underlying molecular mechanism for the observed discrepancy was proposed, where strong activation of EGFR induced EMT in HNSCC cells, and the extracellular domain of EpCAM (EpEX) acted as a novel EGFR ligand that counteracted EGF-induced EMT (Pan *et al.*, 2018). Therefore, dynamic expression of EpCAM is highly involved in EMT during pathological processes, e.g. during cancer progression.

### **5.7 RIP as a potential regulatory mechanism of the cell surface expression of EpCAM**

The accurate timing of the differential regulation of EpCAM in the 3D-model of ESCs differentiation has been addressed in previous sections in the discussion. The analysis of the expression of EpCAM in kinetics of EBs throughout spontaneous differentiation disclosed that EpCAM is completely lost in cellular subsets within a time frame of 12 h, typically between 3.5 and 4 days of differentiation (Sarrach *et al.*, 2018). Interestingly, the down-regulation of transcripts levels of the *EPCAM* gene was slightly delayed compared to the abovementioned timing. In combination with a reported half-life of the EpCAM protein in cancer cells of 21 h (Munz *et al.*, 2008), these data suggested that a post-translational

mechanism might be instrumental to shut down EpCAM expression within a short timeframe.



**Fig. 34: Schematic illustration of the regulated intramembrane proteolysis (RIP) and degradation of EpCAM.**

The Cleavage of EpCAM via RIP, including ADAM10/17 and  $\gamma$ -secretase involved. The cleavage products with degradation efficiency and half-lives of EpCTF in different cell lines are presented on scheme.

Endocytosis and RIP act as two main mechanisms for cells to withdraw mouse and human EpCAM from the cell surface and to subsequently degrade it (**Figure 34**) (Hachmeister *et al.*, 2013; Tsaktanis *et al.*, 2015). Additionally, RIP constitutes the core of the molecular mechanism of signaling functions of EpCAM in stem cells and cancer cells. EpICD, which is generated through RIP of EpCAM, is instrumental in the activation of genes involved in the regulation of cell cycle, proliferation, fatty acid metabolism, differentiation and pluripotency-associated gene expression as described (Maetzel *et al.*, 2009; Munz, Baeuerle and Gires, 2009; Lu *et al.*, 2010; Lin *et al.*, 2012; Chaves-Pérez *et al.*, 2013; Kuan *et al.*, 2017). On that account, it is important to understand the pace of the EpICD

generation in order to further interpret EpCAM's function(s) in differentiation, as well as cell signaling.

RIP involves the initial cleavage of EpCAM by  $\alpha$ - and/or  $\beta$ -secretases within the extracellular domain to generate a soluble ectodomain EpEX together with an EpCTF fragment, which remains resident in the plasma membrane. This EpCTF is then a substrate for the  $\gamma$ -secretase complex that cleaves it at  $\gamma$  and  $\varepsilon$  sites within the transmembrane domain. Cleavage at  $\gamma$ -sites of mouse and human EpCTF releases a small extracellular fragment, which owing to its localization within EpCAM and to the mode of generation was termed A $\beta$ -like fragment, in analogy to the amyloidogenic A $\beta$ -fragment of amyloid precursor protein APP in Alzheimer's syndrome (Hachmeister *et al.*, 2013; Tsaktanis *et al.*, 2015). By now, little is known about the frequency of the first cleavage of EpCAM. A requirement for cell-cell contact to induce the initial cleavage was postulated and soluble EpEX was shown to be a ligand for full length EpCAM (Denzel *et al.*, 2009; Maetzel *et al.*, 2009; Hachmeister *et al.*, 2013). It is also conceivable that a so far unknown soluble or membrane-tethered ligand is required for the induction of EpEX shedding by  $\alpha$ - and/or  $\beta$ -secretases. However, such a ligand has not been identified and thus remains elusive.

EpCAM regulation on cell surface is strongly associated with RIP, repression of RIP by using  $\beta$ -secretase inhibition C3 lead to retention of full-length EpCAM in ESCs (Hachmeister *et al.*, 2013). In contrast, membrane assays at pH4 which represents the pH optimum of  $\beta$ -secretase demonstrate a significant decreases of EpCAM and a strong increases of EpCTF (Hachmeister *et al.*, 2013). Similar results were also confirmed in human EpCAM with head and neck tumour cells (Tsaktanis *et al.*, 2015). Recently, EGF/EGFR-triggered activation of RIP of EpCAM at the cell membrane has been reported, in which EGF treatment of an endometrial carcinoma line RL95-2 could induce the loss of EpEX at the cell membrane (Hsu *et al.*, 2016). Unexpectedly, treatment of RL95-2 cells with EGF in combination with a  $\gamma$ -secretase inhibitor reverted EGF-mediated cleavage of

EpCAM and resulted in the retention of full-length EpCAM at the plasma membrane. However, those findings are in contradiction with our own published data, where EpCAM cleavage was not induced by EGF treatment and where the  $\gamma$ -secretase complex catalyzes EpCAM CTF, not full length EpCAM (Maetzel *et al.*, 2009; Hachmeister *et al.*, 2013; Tsaktanis *et al.*, 2015). In fact, ectodomain shedding by the  $\gamma$ -secretase complex has, to the best of our knowledge, never been reported and generally contradicts the sequential cleavage during RIP. Therefore, it appears hardly conceivable that inhibition of  $\gamma$ -secretase has any direct effect on EGF/EGFR-mediated shedding of EpEX.

Following the first cleavage of EpCAM, the signaling pace as well as the full degradation of EpCAM will be decided by the speed and rate of EpCTF proteolysis by  $\gamma$ -secretase, to release EpICD from plasma membrane. Thereafter, regulation of the stability of EpICD in cells is a means to control the extent and length of nuclear signaling and or other functions of EpCAM. Thus, the efficiency and pace of EpCTF cleavage by  $\gamma$ -secretase was addressed in-depth in the present thesis.

### **5.8 Proteolysis of EpCTF by $\gamma$ -secretase is a slow process**

Using biochemical approaches together with fluorescent protein-tagged variants of mouse and human EpCTF helped to trace and quantify EpCTF cleavage over time. This demonstrated that EpCTF is characterized by a 50% protein turnover in the time range of 45 min to 5.3 h in various cell lines and for both, mouse and human EpCTF (**Figure 34**). Hence cleavage of EpCTF is a particularly slow process as compared to other classical enzymatic processes, such as for example 0.2083 s for the Renin-Angiotensinogen system and 0.0053 s for Chymosin k-Casein reaction (Vreeman *et al.*, 1986; Nguyen *et al.*, 2002). Similar findings were reported for the amyloid precursor protein (APP) CTF, which showed a 50% protein turnover of 2.9 h mediated by  $\gamma$ -secretase *in vitro* (Kamp *et al.*, 2015), suggesting a consistently slow cleavage pace of the  $\gamma$ -secretase complex.

According to the swapping experiments performed in the present thesis, the provenance of EpCTF with respect to species did not affect proteolysis pace. Actually, proteolysis pace of EpCTF variants transfected into cells was dictated by cell lines, but not by the species of origin of the substrates. This was demonstrated by two major findings. Firstly, mouse and human EpCTFs had different half-lives depending on the cell lines in which they were expressed. Secondly, mouse and human EpCTF half-lives did not differ, when expressed in the same cell line. In order to investigate the molecular basis of these findings, evaluation of the aa sequences of mouse and human EpCAM was performed. The entire mouse EpCTF is 91% identical to human EpCTF (**Figure 30**). Importantly, the similarity between mouse and human EpCTF transmembrane domains, where the  $\gamma$  and  $\varepsilon$  cleavages take place, is as high as 96% (**Figure 30**), demonstrating a high similarity of aa sequences of mouse and human EpCTF, which could account for the molecular basis of similar half-lives dictated by cell lines.

Moreover, endogenous EpCTF had a similar slow cleavage pace in carcinoma cells compared with exogenously expressed EpCTF variants, with 4.7 and 5.5 h, respectively (**Figure 34**). Therefore, our experimental system of exogenous EpCTF variants is an appropriate model to reflect the endogenous EpCTF cleavage process.

In conclusion, these findings therefore strongly support the notion that  $\gamma$ -secretase represents the rate-limiting step rather than its substrate EpCTF.

### **5.9 Sustained activation of EpCAM signaling *via* RIP and efficient shutdown by the proteasome**

The release of EpICD from the membrane-tethered EpCTF by  $\gamma$ -secretase is the final rate-limiting step that is required to trigger intracellular signaling events mediated by EpCAM (Münz *et al.*, 2004; Münz, Zeidler and Gires, 2005; Maetzel *et al.*, 2009; Munz, Baeuerle and Gires, 2009; Chaves-Pérez *et al.*, 2013). Given that EpCTF cleavage by  $\gamma$ -secretase is a particularly slow process, signaling by EpCAM through RIP is very unlikely to allow for

a swift reaction to extracellular cues, as is reported for receptors such as EGFR and others (Weng *et al.*, 2004; Oda *et al.*, 2005; Yarden and Shilo, 2007; Andersson, Sandberg and Lendahl, 2011). Therefore, the data presented in this thesis speak in favor of a more stable, steady signaling relying on EplCD in cancer cells (Maetzel *et al.*, 2009; Chaves-Pérez *et al.*, 2013) and in stem cells (Lu *et al.*, 2010; Huang *et al.*, 2011), rather than in favor of a rapid transmission of extracellular cues into cells to transiently active downstream genes. Additionally, the efficiency of EplCD degradation via the proteasome was as high as 94%, which could allow to regulate the strength of EplCD signals and, eventually, to strictly shut down the signaling mediated by EplCD. In addition, differences between malignant cells and normal cells in terms of EplCD nuclear localization (Maetzel *et al.*, 2009) are likely to affect EpCAM signaling via RIP and may cause a signaling deficiency in normal tissues, which has been recently demonstrated in human liver cells (Gerlach *et al.*, 2018). Taken all above, proteolysis of EpCTF by  $\gamma$ -secretase is a slow process and degradation of EplCD by the proteasome is highly efficient.

## 6. SUMMARY

Embryonic development of mammalian species is a highly complex biological process, which is strongly associated with cellular dynamics and morphogenetic mechanisms. However, the current knowledge of precise timing of mouse embryogenesis is still incomplete. Fortunately, genetic manipulation techniques and cultivation of pluripotent ESCs *in vitro* could provide new opportunities to fill these knowledge gaps. In the present thesis, we are aiming at analyzing the spatiotemporal regulation of the epithelial cell adhesion molecule EpCAM in ESCs.

In both E14TG2 $\alpha$  and Bruce4 mouse embryonic stem cell lines, high levels of EpCAM expression is found at the pluripotent stage, while EpCAM expression is significantly reduced after 3D spontaneous differentiation, which closely mimics the early mouse embryogenesis. Interestingly, this dynamic of EpCAM expression is strictly related to the three germ layers of embryoid bodies (EBs). The expression of EpCAM in nascent visceral endoderm was confirmed by the results, showing a co-localization of EpCAM with Foxa2 in marginal cells of EBs. Immunofluorescence double-staining of EpCAM and vimentin displayed a mutually exclusive expression pattern in EBs, demonstrating an absence of EpCAM in mesodermal cells. Genetic knockout of EpCAM in pluripotent ESCs inhibited cardiomyocytes formation. Additionally, cardiomyocytes precursors need a physical contact with Gata4-producing Sox17<sup>+</sup>/EpCAM<sup>+</sup> visceral endoderm for the proper differentiation of cardiomyocytes (Pal and Khanna, 2005; Holtzinger, Rosenfeld and Evans, 2010). Therefore, EpCAM is required for the full differentiation of ESCs. Moreover, a differentiation block at a Mesp1<sup>+</sup> stage was determined in absence of EpCAM. ERas, a Ras protein expressed in ESCs, was co-precipitated with EpCAM in ESCs lysates. Knockout of ERas led to further impairment of the contracting ability of EpCAM knockout-derived EBs, while EpCAM and ERas overexpression both inhibited cardiomyocytes

formation. Thus, these findings describe a novel signaling cascade of EpCAM/ERAs/AKT to support the proper differentiation of ESCs, including to cardiomyocytes.

EpCAM is completely lost in mesodermal cells within 12 h during spontaneous differentiation of EBs. Interestingly, the down-regulation of EpCAM at the transcriptional level was slightly delayed compared to protein loss. In addition, the half-life of the EpCAM protein is 21 h (Munz *et al.*, 2008). Thus, these findings suggested a post-translational mechanism that is in place to shut down EpCAM expression within a short timeframe. At the cell membrane, EpCAM could undergo RIP to generate EpEX and EpCTF fragment. The EpCTF fragment will be further cleaved by the  $\gamma$ -secretase complex to generate EpICD, which is important for EpCAM signaling. Therefore, RIP is a potential regulatory post-translational mechanism for EpCAM expression. The protein turnover of EpCTF was assessed using biochemical and time-lapse imaging techniques in combination with fluorescence-tagged versions of human and mouse EpCTF. By doing so, a time range of 45 min to 5.3 h for a 50% turnover was determined in various cell lines, for both mouse and human EpCTF. However, this EpCTF proteolysis by  $\gamma$ -secretase is a particularly slow process compared to other classical enzymatic processes such as 0.2083 s for Renin-Angiotensinogen system. Such a slow enzymatic processes is dictated by the cell line, but not by the substrate EpCTF, as defined by swapping experiments across species and EpCTF variants. Accordingly, a 96% similarity of aa sequence was revealed between mouse and human EpCAM transmembrane domain, where cleavage takes place. Given that the EpCTF cleavage by  $\gamma$ -secretase is a particularly slow process, EpCAM signaling by RIP is very likely to be a sustained, steady signaling, relying on EpICD in cancer cells and ESCs. Moreover, EpICD degradation via proteasome was as high as 94%. Hence, such an efficient degradation could strictly shut down the EpCAM signaling. Thus, EpCAM is slowly cleaved by  $\gamma$ -secretase followed by efficient proteasomal degradation of EpICD to control EpCAM expression and functionality.

## 7. ZUSAMMENFASSUNG

Die Embryonalentwicklung von Säugetieren ist ein hochkomplexer biologischer Prozess, der stark mit der Zelldynamik und den morphogenetischen Mechanismen verbunden ist. Aktuell ist jedoch wenig über das genaue Timing der Mausembryogenese bekannt. Genetische Manipulationstechniken und die Kultivierung pluripotenter ESCs *in vitro*, bieten neue Möglichkeiten diese Wissenslücken zu schließen. In der vorliegenden Arbeit wollen wir die räumlich-zeitliche Regulation des Epithelzelladhäsionsmoleküls EpCAM in ESCs untersuchen.

Sowohl die embryonalen E14-TG2 $\alpha$  als auch die Bruce4 Maus Stammzelllinien weisen eine hohe EpCAM Expression im pluripotenten Stadium auf, während die Expression nach spontaner 3D-Differenzierung signifikant reduziert ist. Dies bildet ein *in vitro* Modell, welches die frühe Mausembryogenese stark nachahmt. Interessanterweise hängt diese Dynamik der EpCAM Expression eng mit der Ausbildung der drei Keimschichten in *embryoid bodies* (EBs) zusammen. Die Expression von EpCAM im entstehenden viszeralen Endoderm wurde durch eine Ko-Lokalisation von EpCAM mit Foxa2 in Randzellen von EBs gezeigt. Eine Doppelfärbung von EpCAM und Vimentin in EBs zeigt ein sich ausschließendes Expressionsmuster, wobei Vimentin-exprimierende mesodermale Zellen kein EpCAM exprimierten. Ein genetischer Knockout von EpCAM in pluripotenten ESCs hemmte die Bildung von Kardiomyozyten. Zusätzlich benötigen Kardiomyozyten-Vorläufer einen physischen Kontakt mit Gata4-produzierendem Sox17<sup>+</sup>/EpCAM<sup>+</sup> Zellen des Viszeralendoderms, um zu Kardiomyozyten heranzureifen (Pal and Khanna, 2005; Holtzinger, Rosenfeld and Evans, 2010). Daher ist EpCAM für die vollständige Differenzierung von ESCs erforderlich. Darüber hinaus wurde in Abwesenheit von EpCAM ein Differenzierungsblock von Kardiomyozyten im Mesp1<sup>+</sup> Stadium bestimmt. ERas, ein in ESCs exprimiertes Ras-Protein, wurde als Interaktionspartner von EpCAM

charakterisiert. Der Knockout von ERas in ESCs führte zu einer weiteren Beeinträchtigung der Kontraktionsfähigkeit von EpCAM Knockout EBs, während die Überexpression von EpCAM und ERas die Bildung von Kardiomyozyten inhibierte. Daher beschreiben diese Ergebnisse eine neuartige Signalkaskade von EpCAM/ERas/AKT, die für eine ordnungsgemäße Differenzierung von ESCs, einschließlich der Bildung von Kardiomyozyten, benötigt wird.

Während der spontanen Differenzierung von EBs geht EpCAM in mesodermalen Zellen innerhalb von 12 Stunden vollständig verloren. Interessanterweise war die Herunterregulierung von EpCAM auf Transkriptionsebene im Vergleich zum Proteinverlust etwas verzögert. Außerdem beträgt die Halbwertszeit des EpCAM-Proteins 21 Stunden (Munz *et al.*, 2008). Diese Ergebnisse legen daher einen posttranslationalen Mechanismus nahe, mit dem die EpCAM Expression innerhalb eines kurzen Zeitraums eingestellt werden kann. An der Zellmembran könnte EpCAM durch regulierte Intramembranproteolyse (RIP) gespalten werden, um ein EpEX (EpCAM extracellular domain) und ein EpCTF (EpCAM C-terminal fragment) zu erzeugen. Das entstandene EpCTF Fragment wird durch den  $\gamma$ -Sekretasenkomplex weiter gespalten, um EpICD (EpCAM intracellular domain) zu erzeugen, was für die EpCAM Signalübertragung wichtig ist. Daher ist RIP ein potenzieller regulatorischer posttranslationaler Mechanismus der EpCAM-Regulierung. Der Proteinumsatz von EpCTF wurde unter Verwendung von Zeitraffer Bildgebungstechniken an Lebendzellen in Kombination mit fluoreszenzmarkierten Versionen von menschlichem und mousem EpCTF gemessen. Auf diese Weise wurde die Proteinspaltung (50% Umsatz) gemessen, welche in verschiedenen Zelllinien, sowohl für Mäuse als auch für humanes EpCAM, 45 Min. bis 5.3 Std. betrug. Diese EpCTF-Proteolyse durch die  $\gamma$ -Sekretase ist jedoch ein besonders langsamer Prozess. Andere klassische enzymatische Prozesse, wie z.B. das Renin-Angiotensinogen-System benötigen lediglich 0.2083 Sekunden. Die langsame enzymatische Spaltung von EpCTF wird von der  $\gamma$ -Sekretase bestimmt, nicht aber von

dem Substrat EpCTF. Dementsprechend wurde eine 96%-ige Homologie in der Sequenz zwischen der mousen und der humanen EpCAM Transmembrandomäne festgestellt, in der die Spaltung stattfindet. Da die EpCTF-Spaltung durch die  $\gamma$ -Sekretase ein besonders langsamer Prozess ist, ist es sehr wahrscheinlich, dass die EpCAM-Signalübertragung durch RIP eine anhaltende Signalübertragung darstellt, die in Krebszellen und ESCs auf EpICD beruht. Darüber hinaus lag der EpICD Abbau über das Proteasom bei 94%. Daher könnte eine solche effiziente Degradierung die EpCAM Signalwege streng regulieren. Zusammenfassend wird EpCAM langsam durch die  $\gamma$ -Sekretase gespalten, gefolgt von einem effizienten proteasomalen Abbau von EpICD, welcher die Expression und Funktionalität von EpCAM kontrolliert.

# 8. APPENDIX

## ABBREVIATIONS

°C	degree Celsius
aa	amino acids
APS	ammonium persulfate
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
DMEM	Dulbecco`s Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ECL	enhanced chemiluminescence
EDTA	ethylene diamine tetra acetic acid
EMT	epithelial to mesenchymal transition
EpCAM	epithelial cell adhesion molecule
EpCTF	C-terminal fragment of EpCAM
EpICD	intracellular domain of EpCAM
GFP	green fluorescent protein
FACS	fluorescence activated cell sorting
FC	flow cytometry
FCS	fetal calf serum
IH	immunohistochemistry
IF	immunofluorescence
KCl	potassium chloride
kDa	kilo Dalton
L	liter
M	molar
mA	milli ampere
max	maximal
mg	milligram
µg	microgram
ng	Nanogram
min	minute

mL	milliliter
$\mu$ L	microliter
mM	millimolar
$\mu$ M	micromolar
nM	Nanomolar
mRNA	messenger RNA
NaCl	sodium chloride
N-term	N-terminus
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBST	PBS + Tween-20
PCR	polymerase chain reaction
PFA	paraformaldehyde
PVDF	polyvinylidene fluoride
qRT-PCR	quantitative Real Time PCR
rcf	relative centrifugal force
RIP	regulated intramembrane proteolysis
RNA	ribonucleic acid
rpm	revolutions per minute
RT	reverse transcriptase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide
SILAC	stable isotope labeling by/with amino acids in cell
TEMED	Tetramethylenediamine
TRIS	tris (hydroxyl methyl) aminomethane
Triton X-100	polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl
WB	western blot
w/o	without
YFP	yellow fluorescent protein
$\alpha$	alpha
$\beta$	beta

## 9. REFERENCE

Aach, J. *et al.* (2017) “Addressing the ethical issues raised by synthetic human entities with embryo-like features,” *eLife*. doi: 10.7554/eLife.20674.

Adewumi, O. *et al.* (2007) “Characterization of human embryonic stem cell lines by the International Stem Cell Initiative,” *Nature Biotechnology*. doi: 10.1038/nbt1318.

Andersson, E. R., Sandberg, R. and Lendahl, U. (2011) “Notch signaling: Simplicity in design, versatility in function,” *Development*. doi: 10.1242/dev.063610.

Assou, S. *et al.* (2007) “A Meta-Analysis of Human Embryonic Stem Cells Transcriptome Integrated into a Web-Based Expression Atlas,” *Stem Cells*. doi: 10.1634/stemcells.2006-0352.

BAI, H. *et al.* (2013) “Expression and Potential Role of GATA Factors in Trophoblast Development,” *Journal of Reproduction and Development*, 59(1), pp. 1–6. doi: 10.1262/JRD.2012-100.

Balzar, M. *et al.* (1998) “Cytoplasmic tail regulates the intercellular adhesion function of the epithelial cell adhesion molecule.,” *Molecular and cellular biology*, 18(8), pp. 4833–43. doi: 10.1073/pnas.76.3.1438.

Balzar, M. *et al.* (1999a) “The biology of the 17-1A antigen (Ep-CAM).,” *Journal of molecular medicine (Berlin, Germany)*, 77, pp. 699–712. doi: 10.1007/s001099900038.

Baumeister, P. *et al.* (2018) “High Expression of EpCAM and Sox2 is a Positive Prognosticator of Clinical Outcome for Head and Neck Carcinoma,” *Scientific Reports*. doi: 10.1038/s41598-018-32178-8.

Bedell, M. A. *et al.* (1997) “Mouse models of human disease. Part II: Recent progress and future directions,” *Genes and Development*. doi: 10.1101/gad.11.1.11.

Bedell, M. A., Jenkins, N. A. and Copeland, N. G. (1997) “Mouse models of human disease. Part I: Techniques and resources for genetic analysis in mice,” *Genes and Development*. doi: 10.1101/gad.11.1.1.

Bedzhov, I. *et al.* (2014) “Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo,” *Philosophical Transactions of the Royal*

*Society B: Biological Sciences*. doi: 10.1098/rstb.2013.0538.

ten Berge, D. *et al.* (2008) "Wnt Signaling Mediates Self-Organization and Axis Formation in Embryoid Bodies," *Cell Stem Cell*. doi: 10.1016/j.stem.2008.09.013.

Bergsagel, P. L. *et al.* (1992) "A murine cDNA encodes a pan-epithelial glycoprotein that is also expressed on plasma cells," *J Immunol*.

Beyer Nardi, N. and da Silva Meirelles, L. (2006) "Mesenchymal stem cells: isolation, in vitro expansion and characterization.," *Handbook of experimental pharmacology*.

Bhattacharya, B. *et al.* (2004) "Gene expression in human embryonic stem cell lines: Unique molecular signature," *Blood*. doi: 10.1182/blood-2003-09-3314.

Blanpain, C. and Simons, B. D. (2013) "Unravelling stem cell dynamics by lineage tracing," *Nature Reviews Molecular Cell Biology*, 14(8), pp. 489–502. doi: 10.1038/nrm3625.

Boheler, K. R. *et al.* (2002) "Differentiation of pluripotent embryonic stem cells into cardiomyocytes," *Circulation Research*, pp. 189–201. doi: 10.1161/01.RES.0000027865.61704.32.

Bondue, A. *et al.* (2008) "Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification," *Cell Stem Cell*. doi: 10.1016/j.stem.2008.06.009.

Boyer, L. A. *et al.* (2005) "Core transcriptional regulatory circuitry in human embryonic stem cells," *Cell*. doi: 10.1016/j.cell.2005.08.020.

Burtscher, I. and Lickert, H. (2009a) "Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo," *Development*. doi: 10.1242/dev.028415.

Burtscher, I. and Lickert, H. (2009b) "Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo," *Development*. doi: 10.1242/dev.028415.

Buttery, L. D. K. *et al.* (2001) "Differentiation of osteoblasts and in Vitro bone formation from murine embryonic stem cells," *Tissue Engineering*. doi: 10.1089/107632700300003323.

Chadeneau, C. *et al.* (1991) "Characterization, isolation and amino terminal sequencing of a rat colon carcinoma-associated antigen," *International Journal of Cancer*. doi: 10.1002/ijc.2910470620.

- Chaves-Pérez, A. *et al.* (2013) “EpCAM regulates cell cycle progression via control of cyclin D1 expression,” *Oncogene*. doi: 10.1038/onc.2012.75.
- Chen, M. *et al.* (2015) “Enrichment of cardiac differentiation by a large starting number of embryonic stem cells in embryoid bodies is mediated by the Wnt11-JNK pathway,” *Biotechnology Letters*. doi: 10.1007/s10529-014-1700-5.
- Chen, X. *et al.* (2008) “Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells,” *Cell*. doi: 10.1016/j.cell.2008.04.043.
- Denzel, S. *et al.* (2009) “Initial activation of EpCAM cleavage via cell-to-cell contact,” *BMC Cancer*. doi: 10.1186/1471-2407-9-402.
- Desbaillets, I. *et al.* (2000) “Embryoid bodies: An in vitro model of mouse embryogenesis,” *Experimental Physiology*. doi: 10.1111/j.1469-445X.2000.02104.x.
- Doetschman, T. C. *et al.* (1985) “The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium.,” *Journal of embryology and experimental morphology*.
- Evans, M. J. and Kaufman, M. H. (1981) “Establishment in culture of pluripotential cells from mouse embryos,” *Nature*. doi: 10.1038/292154a0.
- Foley, A. C. *et al.* (2006) “Embryonic heart induction,” in *Annals of the New York Academy of Sciences*. doi: 10.1196/annals.1380.008.
- Fuchs, E., Tumber, T. and Guasch, G. (2004) “Socializing with the neighbors: Stem cells and their niche,” *Cell*. doi: 10.1016/S0092-8674(04)00255-7.
- Gaber, A. *et al.* (2018) “EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion,” *Scientific Reports*. doi: 10.1038/s41598-018-31482-7.
- Gerlach, J. C. *et al.* (2018) “Epithelial cell adhesion molecule fragments and signaling in primary human liver cells,” *Journal of Cellular Physiology*. doi: 10.1002/jcp.26286.
- Gires, O., Klein, C. A. and Baeuerle, P. A. (2009) “On the abundance of EpCAM on cancer stem cells,” *Nature Reviews Cancer*. doi: 10.1038/nrc2499-c1.
- Gires, O. and Stoecklein, N. H. (2014) “Dynamic EpCAM expression on circulating and disseminating tumor cells: Causes and consequences,” *Cellular and Molecular Life*

*Sciences*. doi: 10.1007/s00018-014-1693-1.

González, B. *et al.* (2009) "EpCAM is involved in maintenance of the murine embryonic stem cell phenotype," *Stem Cells*, 27(8), pp. 1782–1791. doi: 10.1002/stem.97.

Gorges, T. M. *et al.* (2012) "Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition," *BMC Cancer*. doi: 10.1186/1471-2407-12-178.

van der Gun, B. T. F. *et al.* (2010) "EpCAM in carcinogenesis: The good, the bad or the ugly," *Carcinogenesis*. doi: 10.1093/carcin/bgq187.

Hachmeister, M. *et al.* (2013) "Regulated Intramembrane Proteolysis and Degradation of Murine Epithelial Cell Adhesion Molecule mEpCAM," *PLoS ONE*. doi: 10.1371/journal.pone.0071836.

Hayes, M. *et al.* (2012) "Clinical review: Stem cell therapies for acute lung injury/acute respiratory distress syndrome - hope or hype?," *Critical Care*. doi: 10.1186/cc10570.

Herlyn, M. *et al.* (1979) "Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies.," *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.76.3.1438.

Holtzinger, A., Rosenfeld, G. E. and Evans, T. (2010) "Gata4 directs development of cardiac-inducing endoderm from ES cells," *Developmental Biology*. doi: 10.1016/j.ydbio.2009.10.003.

Hsu, Y. T. *et al.* (2016) "EpCAM-regulated transcription exerts influences on nanomechanical properties of endometrial cancer cells that promote epithelial-to-mesenchymal transition," *Cancer Research*. doi: 10.1158/0008-5472.CAN-16-0752.

Huang, H. P. *et al.* (2011) "Epithelial cell adhesion molecule (EpCAM) complex proteins promote transcription factor-mediated pluripotency reprogramming," *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.256164.

Huang, L. *et al.* (2018) "Functions of EpCAM in physiological processes and diseases (Review)," *International Journal of Molecular Medicine*. doi: 10.3892/ijmm.2018.3764.

Hwang, Y.-S. *et al.* (2009) "Microwell-mediated control of embryoid body size regulates embryonic stem cell fate via differential expression of WNT5a and WNT11," *Proceedings*

of the National Academy of Sciences. doi: 10.1073/pnas.0905550106.

Itskovitz-Eldor, J. *et al.* (2000) "Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers.," *Molecular medicine (Cambridge, Mass.)*. doi: 10859025.

Jaenisch, R. and Young, R. (2008) "Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming," *Cell*. doi: 10.1016/j.cell.2008.01.015.

Johnson, M. H. (2009) "From Mouse Egg to Mouse Embryo: Polarities, Axes, and Tissues," *Annual Review of Cell and Developmental Biology*. doi: 10.1146/annurev.cellbio.042308.113348.

Kamimoto, K. *et al.* (2016) "Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling," *eLife*. doi: 10.7554/eLife.15034.

Kamp, F. *et al.* (2015) "Intramembrane proteolysis of  $\beta$ -amyloid precursor protein by  $\gamma$ -secretase is an unusually slow process," *Biophysical Journal*. doi: 10.1016/j.bpj.2014.12.045.

Kempers, M. J. E. *et al.* (2011) "Risk of colorectal and endometrial cancers in EPCAM deletion-positive Lynch syndrome: A cohort study," *The Lancet Oncology*. doi: 10.1016/S1470-2045(10)70265-5.

Knowles, B. B., Aden, D. P. and Solter, D. (1978) "Monoclonal antibody detecting a stage-specific embryonic antigen (SSEA-1) on preimplantation mouse embryos and teratocarcinoma cells.," *Current Topics in Microbiology and Immunology*.

Kojima, Y., Tam, O. H. and Tam, P. P. L. (2014) "Timing of developmental events in the early mouse embryo," *Seminars in Cell and Developmental Biology*. doi: 10.1016/j.semcdb.2014.06.010.

Kuan, I. I. *et al.* (2017) "EpEX/EpCAM and Oct4 or Klf4 alone are sufficient to generate induced pluripotent stem cells through STAT3 and HIF2 $\alpha$ ," *Scientific Reports*. doi: 10.1038/srep41852.

Kuiper, R. P. *et al.* (2011) "Recurrence and variability of germline EPCAM deletions in Lynch syndrome," *Human Mutation*. doi: 10.1002/humu.21446.

- Lal, M. and Caplan, M. (2011) "Regulated Intramembrane Proteolysis: Signaling Pathways and Biological Functions," *Physiology*. doi: 10.1152/physiol.00028.2010.
- Lee, M. Y. *et al.* (2009a) "Interaction of galectin-1 with caveolae induces mouse embryonic stem cell proliferation through the Src, ERas, Akt and mTOR signaling pathways," *Cellular and Molecular Life Sciences*. doi: 10.1007/s00018-009-8691-8.
- Lee, M. Y. *et al.* (2009b) "Interaction of galectin-1 with caveolae induces mouse embryonic stem cell proliferation through the Src, ERas, Akt and mTOR signaling pathways," *Cellular and Molecular Life Sciences*. doi: 10.1007/s00018-009-8691-8.
- Lei, Z. *et al.* (2012) "EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins," *Developmental Biology*. doi: 10.1016/j.ydbio.2012.07.005.
- Lian, Q. *et al.* (2007) "Derivation of Clinically Compliant MSCs from CD105+, CD24-Differentiated Human ESCs," *Stem Cells*. doi: 10.1634/stemcells.2006-0420.
- Liang, K. H. *et al.* (2018) "Extracellular domain of EpCAM enhances tumor progression through EGFR signaling in colon cancer cells," *Cancer Letters*. doi: 10.1016/j.canlet.2018.06.040.
- Ligtenberg, M. J. L. *et al.* (2009) "Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1," *Nature Genetics*. doi: 10.1038/ng.283.
- Lin, C. W. *et al.* (2012) "Epithelial cell adhesion molecule regulates tumor initiation and tumorigenesis via activating reprogramming factors and epithelial-mesenchymal transition gene expression in colon cancer," *Journal of Biological Chemistry*. doi: 10.1074/jbc.M112.386235.
- Litvinov, S. V., Velders, M. P., *et al.* (1994) "Ep-CAM: A human epithelial antigen is a homophilic cell-cell adhesion molecule," *Journal of Cell Biology*. doi: 10.1083/jcb.125.2.437.
- Litvinov, S. V., Bakker, H. A. M., *et al.* (1994) "Evidence for a role of the epithelial glycoprotein 40 (ep-CAM) in epithelial cell-cell adhesion," *Cell Communication and Adhesion*. doi: 10.3109/15419069409004452.
- Litvinov, S. V. *et al.* (1997) "Epithelial cell adhesion molecule (Ep-CAM) modulates cell-cell

interactions mediated by classic cadherins,” *Journal of Cell Biology*. doi: 10.1083/jcb.139.5.1337.

Liu, J. *et al.* (2009) “Integrins are required for the differentiation of visceral endoderm,” *Journal of Cell Science*. doi: 10.1242/jcs.037663.

Liu, X. *et al.* (2019) “Epithelial-type systemic breast carcinoma cells with a restricted mesenchymal transition are a major source of metastasis,” *Science Advances*. doi: 10.1126/sciadv.aav4275.

Loh, Y. H. *et al.* (2006) “The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells,” *Nature Genetics*. doi: 10.1038/ng1760.

Lu, H. *et al.* (2013) “EpCAM Is an Endoderm-Specific Wnt Derepressor that Licenses Hepatic Development,” *Developmental Cell*. doi: 10.1016/j.devcel.2013.01.021.

Lu, T. Y. *et al.* (2010) “Epithelial cell adhesion molecule regulation is associated with the maintenance of the undifferentiated phenotype of human embryonic stem cells,” *Journal of Biological Chemistry*. doi: 10.1074/jbc.M109.077081.

Maetzel, D. *et al.* (2009) “Nuclear signalling by tumour-associated antigen EpCAM,” *Nature Cell Biology*. doi: 10.1038/ncb1824.

Maimets, M. *et al.* (2016) “Long-Term in Vitro Expansion of Salivary Gland Stem Cells Driven by Wnt Signals,” *Stem Cell Reports*. doi: 10.1016/j.stemcr.2015.11.009.

Martin, G. R. (1981) “Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.,” *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.78.12.7634.

Mazzotta, S. *et al.* (2016a) “Distinctive Roles of Canonical and Noncanonical Wnt Signaling in Human Embryonic Cardiomyocyte Development,” *Stem Cell Reports*. doi: 10.1016/j.stemcr.2016.08.008.

Mazzotta, S. *et al.* (2016b) “Distinctive Roles of Canonical and Noncanonical Wnt Signaling in Human Embryonic Cardiomyocyte Development,” *Stem Cell Reports*. doi: 10.1016/j.stemcr.2016.08.008.

Medina, M. and Dotti, C. G. (2003) “RIPped out by presenilin-dependent  $\gamma$ -secretase,” *Cellular Signalling*. doi: 10.1016/S0898-6568(03)00041-X.

Migeotte, I. *et al.* (2010) "Rac1-dependent collective cell migration is required for specification of the anterior-posterior body axis of the mouse," *PLoS Biology*. doi: 10.1371/journal.pbio.1000442.

Mikawa, T. *et al.* (2004) "Induction and Patterning of the Primitive Streak, an Organizing Center of Gastrulation in the Amniote," *Developmental Dynamics*. doi: 10.1002/dvdy.10458.

Moore, K. A. and Lemischka, I. R. (2006) "Stem cells and their niches," *Science*. doi: 10.1126/science.11110542.

Morrison, S. J. and Spradling, A. C. (2008) "Stem Cells and Niches: Mechanisms That Promote Stem Cell Maintenance throughout Life," *Cell*. doi: 10.1016/j.cell.2008.01.038.

Munz, M. *et al.* (2008) "Glycosylation is crucial for stability of tumour and cancer stem cell antigen EpCAM," *Frontiers in Bioscience*. doi: 3075 [pii].

Münz, M. *et al.* (2004) "The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation," *Oncogene*. doi: 10.1038/sj.onc.1207610.

Munz, M., Baeuerle, P. A. and Gires, O. (2009) "The emerging role of EpCAM in cancer and stem cell signaling," *Cancer Research*. doi: 10.1158/0008-5472.CAN-09-0654.

Münz, M., Zeidler, R. and Gires, O. (2005) "The tumour-associated antigen EpCAM upregulates the fatty acid binding protein E-FABP," *Cancer Letters*. doi: 10.1016/j.canlet.2004.11.048.

Nagy, A. *et al.* (1990) "Embryonic stem cells alone are able to support fetal development in the mouse.," *Development (Cambridge, England)*.

Ng, V. Y. *et al.* (2010) "Characterization of epithelial cell adhesion molecule as a surface marker on undifferentiated human embryonic stem cells," *Stem Cells*. doi: 10.1002/stem.221.

Nguyen, G. *et al.* (2002) "Pivotal role of the renin/prorenin receptor in angiotensin II production and cellular responses to renin," *Journal of Clinical Investigation*. doi: 10.1172/JCI0214276.

Nishikawa, S. I., Jakt, L. M. and Era, T. (2007) "Embryonic stem-cell culture as a tool for developmental cell biology," *Nature Reviews Molecular Cell Biology*. doi:

10.1038/nrm2189.

Nowotschin, S. and Hadjantonakis, A. K. (2010) "Cellular dynamics in the early mouse embryo: From axis formation to gastrulation," *Current Opinion in Genetics and Development*. doi: 10.1016/j.gde.2010.05.008.

Nusse, R. *et al.* (2008) "Wnt signaling and stem cell control," in *Cold Spring Harbor Symposia on Quantitative Biology*. doi: 10.1101/sqb.2008.73.035.

Oda, K. *et al.* (2005) "A comprehensive pathway map of epidermal growth factor receptor signaling," *Molecular Systems Biology*. doi: 10.1038/msb4100014.

Osta, W. A. *et al.* (2004) "EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy," *Cancer Research*. doi: 10.1158/0008-5472.CAN-04-0754.

Paige, S. L. *et al.* (2015) "Molecular regulation of cardiomyocyte differentiation," *Circulation Research*. doi: 10.1161/CIRCRESAHA.116.302752.

Pal, R. and Khanna, A. (2005) "Role of Hepatocyte-like Cells in the Differentiation of Cardiomyocytes from Mouse Embryonic Stem Cells," *Stem Cells and Development*. doi: 10.1089/scd.2005.14.153.

Pan, M. *et al.* (2018) "EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers," *PLoS Biology*. doi: 10.1371/journal.pbio.2006624.

Parameswaran, M. and Tam, P. P. L. (1995) "Regionalisation of cell fate and morphogenetic movement of the mesoderm during mouse gastrulation," *Developmental Genetics*. doi: 10.1002/dvg.1020170104.

Pavšič, M. and Lenarčič, B. (2011) "Expression, crystallization and preliminary X-ray characterization of the human epithelial cell-adhesion molecule ectodomain," *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. doi: 10.1107/S1744309111031897.

Paz, A. *et al.* (2001) "Glectin-1 binds oncogenic H-Ras to mediate Ras membrane anchorage and cell transformation," *Oncogene*. doi: 10.1038/sj.onc.1204950.

Pera, M. F. (2017) "Human embryo research and the 14-day rule," *Development*. doi:

10.1242/dev.151191.

Pessina, A. and Gribaldo, L. (2006) "The key role of adult stem cells: therapeutic perspectives," *Current Medical Research and Opinion*. doi: 10.1185/030079906x148517.

Puram, S. V. *et al.* (2017) "Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer," *Cell*. doi: 10.1016/j.cell.2017.10.044.

Rashidi, H. *et al.* (2018) "3D human liver tissue from pluripotent stem cells displays stable phenotype in vitro and supports compromised liver function in vivo," *Archives of Toxicology*. doi: 10.1007/s00204-018-2280-2.

Rivera-Pérez, J. A., Mager, J. and Magnuson, T. (2003) "Dynamic morphogenetic events characterize the mouse visceral endoderm," *Developmental Biology*. doi: 10.1016/S0012-1606(03)00302-6.

Rizzino, A. (2013) "Concise review: The Sox2-Oct4 connection: Critical players in a much larger interdependent network integrated at multiple levels," *Stem Cells*. doi: 10.1002/stem.1352.

Rodda, D. J. *et al.* (2005) "Transcriptional regulation of Nanog by OCT4 and SOX2," *Journal of Biological Chemistry*. doi: 10.1074/jbc.M502573200.

Rodriguez-Viciano, P. *et al.* (1994) "Phosphatidylinositol-3-OH kinase direct target of Ras," *Nature*. doi: 10.1038/370527a0.

Rumilla, K. *et al.* (2011) "Frequency of deletions of EPCAM (TACSTD1) in MSH2-associated Lynch syndrome cases," *Journal of Molecular Diagnostics*. doi: 10.1016/j.jmoldx.2010.11.011.

Saiz, N. and Plusa, B. (2013) "Early cell fate decisions in the mouse embryo," *Reproduction*, 145(3), pp. R65–R80. doi: 10.1530/REP-12-0381.

Sarrach, S. *et al.* (2018) "Spatiotemporal patterning of EpCAM is important for murine embryonic endo-And mesodermal differentiation," *Scientific Reports*. doi: 10.1038/s41598-018-20131-8.

Schmelzer, E. (2007) "EpCAM expression in normal, non-pathological tissues," *Frontiers in Bioscience*. doi: 10.2741/2911.

- Schmelzer, E. *et al.* (2007) "Human hepatic stem cells from fetal and postnatal donors," *The Journal of Experimental Medicine*. doi: 10.1084/jem.20061603.
- Schnell, U., Cirulli, V. and Giepmans, B. N. G. (2013) "EpCAM: Structure and function in health and disease," *Biochimica et Biophysica Acta - Biomembranes*. doi: 10.1016/j.bbamem.2013.04.018.
- Schön, M. P. *et al.* (1993) "Biochemical and immunological characterization of the human carcinoma - associated antigen MH 99/KS 1/4," *International Journal of Cancer*. doi: 10.1002/ijc.2910550619.
- Shamblott, M. J. *et al.* (1998) "Derivation of pluripotent stem cells from cultured human primordial germ cells," *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.95.23.13726.
- Sivagnanam, M. *et al.* (2008) "Identification of EpCAM as the Gene for Congenital Tufting Enteropathy," *Gastroenterology*. doi: 10.1053/j.gastro.2008.05.036.
- Sivagnanam, M. *et al.* (2010) "Further evidence for EpCAM as the gene for congenital tufting enteropathy," *American Journal of Medical Genetics, Part A*. doi: 10.1002/ajmg.a.33186.
- Skottman, H. *et al.* (2005) "Gene Expression Signatures of Seven Individual Human Embryonic Stem Cell Lines," *Stem Cells*. doi: 10.1634/stemcells.2004-0341.
- Smith, A. G. *et al.* (1988) "Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides," *Nature*. doi: 10.1038/336688a0.
- Smith, L. J. (1980) "Embryonic axis orientation in the mouse and its correlation with blastocyst relationships to the uterus. Part 1. Relationships between 82 hours and 4 1/4 days.," *Journal of embryology and experimental morphology*.
- Solanas, G. *et al.* (2011) "Cleavage of E-cadherin by ADAM10 mediates epithelial cell sorting downstream of EphB signalling," *Nature Cell Biology*. doi: 10.1038/ncb2298.
- Solter, D. and Knowles, B. B. (1978) "Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1).," *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.75.11.5565.
- Sozen, B. *et al.* (2018) "Self-assembly of embryonic and two extra-embryonic stem cell

- types into gastrulating embryo-like structures,” *Nature Cell Biology*. doi: 10.1038/s41556-018-0147-7.
- Spater, D. *et al.* (2014) “How to make a cardiomyocyte,” *Development*. doi: 10.1242/dev.091538.
- Szala, S. *et al.* (1990) “Molecular cloning of cDNA for the carcinoma-associated antigen GA733-2.,” *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.87.9.3542.
- Taft, R. A. (2008) “Virtues and limitations of the preimplantation mouse embryo as a model system,” *Theriogenology*. doi: 10.1016/j.theriogenology.2007.09.032.
- Takahashi, K., Mitsui, K. and Yamanaka, S. (2003) “Role of ERas in promoting tumour-like properties in mouse embryonic stem cells,” *Nature*. doi: 10.1038/nature01646.
- Takaoka, K. and Hamada, H. (2012) “Cell fate decisions and axis determination in the early mouse embryo.,” *Development (Cambridge, England)*, 139(1), pp. 3–14. doi: 10.1242/dev.060095.
- Tam, P. P. L., Williams, E. A. and Chan, W. Y. (1993) “Gastrulation in the mouse embryo: Ultrastructural and molecular aspects of germ layer morphogenesis,” *Microscopy Research and Technique*. doi: 10.1002/jemt.1070260405.
- Thampoe, I. J., Ng, J. S. C. and Lloyd, K. O. (1988) “Biochemical analysis of a human epithelial surface antigen: Differential cell expression and processing,” *Archives of Biochemistry and Biophysics*. doi: 10.1016/0003-9861(88)90040-9.
- Thomson, M. *et al.* (2011) “Pluripotency factors in embryonic stem cells regulate differentiation into germ layers,” *Cell*. doi: 10.1016/j.cell.2011.05.017.
- Tsaktanis, T. *et al.* (2015) “Cleavage and cell adhesion properties of human epithelial cell adhesion molecule (HEPCAM),” *Journal of Biological Chemistry*. doi: 10.1074/jbc.M115.662700.
- De Ugarte, D. A. *et al.* (2003) “Comparison of multi-lineage cells from human adipose tissue and bone marrow,” *Cells Tissues Organs*. doi: 10.1159/000071150.
- Ulrich, F. *et al.* (2005) “Wnt11 functions in gastrulation by controlling cell cohesion through Rab5c and E-Cadherin,” *Developmental Cell*. doi: 10.1016/j.devcel.2005.08.011.

Uosaki, H. *et al.* (2012) "Direct Contact with Endoderm-Like Cells Efficiently Induces Cardiac Progenitors from Mouse and Human Pluripotent Stem Cells," *PLoS ONE*. doi: 10.1371/journal.pone.0046413.

Vannier, C. *et al.* (2013) "Zeb1 regulates E-cadherin and Epcam (epithelial cell adhesion molecule) expression to control cell behavior in early zebrafish development," *Journal of Biological Chemistry*. doi: 10.1074/jbc.M113.467787.

Varner, V. D. and Taber, L. A. (2012) "Not just inductive: a crucial mechanical role for the endoderm during heart tube assembly," *Development*. doi: 10.1242/dev.073486.

Vreeman, H. J. *et al.* (1986) "Characterization of bovine kappa-casein fractions and the kinetics of chymosin-induced macropeptide release from carbohydrate-free and carbohydrate-containing fractions determined by high-performance gel-permeation chromatography.," *The Biochemical journal*. doi: 10.1042/bj2400087.

Wang, H. *et al.* (2016) "Circulating and disseminated tumor cells: diagnostic tools and therapeutic targets in motion," *Oncotarget*. doi: 10.18632/oncotarget.12242.

Wang, H. and Dey, S. K. (2006) "Roadmap to embryo implantation: Clues from mouse models," *Nature Reviews Genetics*. doi: 10.1038/nrg1808.

Wang, X. and Yang, P. (2008) "In vitro Differentiation of Mouse Embryonic Stem (mES) Cells Using the Hanging Drop Method," *Journal of Visualized Experiments*. doi: 10.3791/825.

Weng, A. P. *et al.* (2004) "Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia," *Science*. doi: 10.1126/science.1102160.

De Wert, G. and Mummery, C. (2003) "Human embryonic stem cells: Research, ethics and policy," *Human Reproduction*. doi: 10.1093/humrep/deg143.

Williams, R. L. *et al.* (1988) "Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells," *Nature*. doi: 10.1038/336684a0.

Yarden, Y. and Shilo, B. Z. (2007) "SnapShot: EGFR Signaling Pathway," *Cell*. doi: 10.1016/j.cell.2007.11.013.

Yoon, S. M. *et al.* (2011) "Epithelial cell adhesion molecule (EpCAM) marks hepatocytes newly derived from stem/progenitor cells in humans," *Hepatology*. doi: 10.1002/hep.24122.

Yu, T., Ma, Y. and Wang, H. (2017) “EpCAM Intracellular Domain Promotes Porcine Cell Reprogramming by Upregulation of Pluripotent Gene Expression via Beta-catenin Signaling,” *Scientific Reports*. doi: 10.1038/srep46315.

Yu, Y. *et al.* (2014) “Stimulation of somatic cell reprogramming by ERas-Akt-FoxO1 signaling axis,” *Stem Cells*. doi: 10.1002/stem.1447.

Zhao, Z. A. *et al.* (2015) “The roles of ERAS during cell lineage specification of mouse early embryonic development,” *Open Biology*. doi: 10.1098/rsob.150092.

# 10. PUBLICATIONS

During my work, I was able to contribute to the following publications:

## **First authorships:**

**Membrane-associated epithelial cell adhesion molecule is slowly cleaved by  $\gamma$ -secretase prior to efficient proteasomal degradation of its intracellular domain.**

**Huang Y\***, Chanou A\*, Kranz G, Pan M, Kohlbauer V, Ettinger A, Gires O.

J Biol Chem. 2019 Mar 1;294(9):3051-3064. doi: 10.1074

\*Yuanchi Huang, Anna Chanou contributed equally to this work.

**Spatiotemporal patterning of EpCAM is important for murine embryonic endo- and mesodermal differentiation.**

Sarrach S\*, **Huang Y\***, Niedermeyer S, Hachmeister M, Fischer L, Gille S, Pan M, Mack B, Kranz G, Libl D, Merl-Pham J, Hauck SM, Paoluzzi Tomada E, Kieslinger M, Jeremias I, Scialdone A, Gires O.

Sci Rep. 2018 Jan 29;8(1):1801. doi: 10.1038

\*Sannia Sarrach and Yuanchi Huang contributed equally to this work.

**Co-authorships:**

**EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers.**

Pan M, Schinke H, Luxenburger E, Kranz G, Shakhtour J, Libl D, **Huang Y**, Gaber A, Pavšič M, Lenarčič B, Kitz J, Jakob M, Schwenk-Zieger S, Canis M, Hess J, Unger K, Baumeister P, Gires O.

PLoS Biol. 2018 Sep 27;16(9): e2006624. doi: 10.1371

**Epithelial-type systemic breast carcinoma cells with a restricted mesenchymal transition are a major source of metastasis.**

Liu X, Li J, Cadilha BL, Markota A, Voigt C, Huang Z, Lin PP, Wang DD, Dai J, Kranz G, Krandick A, Libl D, Zitzelsberger H, Zagorski I, Braselmann H, Pan M, Zhu S, **Huang Y**, Niedermeyer S, Reichel CA, Uhl B, Briukhovetska D, Suárez J, Kobold S, Gires O, Wang H.

Sci Adv. 2019 Jun 19;5(6): eaav4275. doi: 10.1126

# 11. ACKNOWLEDGEMENTS

First of all, I want to thank my supervisor Prof. Olivier Gires, for giving me the chance to work in his lab and providing me with the interesting and challenging project of my thesis. Thank you very much for your great supervision, for always taking time to discuss and for all the things I could learn from you during the last years.

A special thank you to Anna Chanou, a friendly Greek girl, my project became such efficient with your help. Of course, I also want to thank our cooperation partners in Helmholtz, especially Andreas Ettinger taught us so much knowledge about image processing and microscope. A lot of thanks to Min Pan and Vera Kohlbauer for always help me to provide the most valuable suggestions on everyday. Thank you for all the support, discussions and brainstorm.

I also want to thank my colleagues Elke Luxenburger, Gisela Kranz, Henrik Schinke, Jean-Édouard Margotin, Sabina Schwenk-Zieger, Xiao Liu and all the other people in our research department. I want to thank you for the wonderful time I could spend with you, for all your help and advice, for all the things I could learn from you and for all the fun we had together.

Finally, thanks to my whole family, especially my parents for all their love, help and support during my whole life, I can never reach the current achievement without you, I love you mom and dad!

Thank you!