Molecular insights into early cell fate specification and pancreatic cancer

Cumulative dissertation submitted in partial fulfillment of the requirements for the degree of „Doctor rerum naturalium” (Dr. rer. nat.) to the Faculty of Natural Sciences University Ulm.

Ronan Russell, BSc., MSc.
Kilkenny

2015
Current dean of the Faculty of Natural Sciences: Prof. Dr. Peter Dürre

First supervisor: Prof. Dr. Michael Kühl

Second Supervisor: PD. Dr. Alexander Kleger

Day doctorate awarded: 23.11.2015
1. Table of Contents

2. ABBREVIATIONS .......................................................................................................................... 2

3. PREFACE ...................................................................................................................................... 4

4. SUMMARY ...................................................................................................................................... 6

5. INTRODUCTION .......................................................................................................................... 8

5.1. PLURIPOTENCY IN MICE AND MEN ......................................................................................... 8

5.1.1. Signalling pathways associated with pluripotency ................................................................. 13

5.1.2. Pluripotency associated transcription factor network .......................................................... 15

5.2. T-BOX FACTORS ...................................................................................................................... 18

5.2.1. T-box factor 3 ........................................................................................................................ 19

5.3. PLURIPOTENCY AND CELL LINEAGE COMMITMENT ............................................................ 21

5.3.1. Gastrulation and Endoderm specification ............................................................................. 21

5.3.2. Spontaneous Differentiation ............................................................................................... 23

5.3.3. Directed Differentiation .................................................................................................. 23

5.4. PANCREATIC DEVELOPMENT AND REGULATING KEY FACTORS ..................................... 26

5.5. PANCREAS ANATOMY AND PHYSIOLOGY ........................................................................... 29

5.5.1. The exocrine pancreas ......................................................................................................... 30

5.5.2. The endocrine pancreas ...................................................................................................... 31

5.6. PANCREATIC DUCTAL ADENOCARCINOMA (PDAC) ............................................................ 31

5.6.1. Disease Pathology ............................................................................................................. 33

5.6.2. Common mutations associated with PDAC .......................................................................... 37

5.6.2.1. Kirsten rat sarcoma viral oncogene homolog (KRAS) ...................................................... 38

5.6.2.2. The TP53 tumor suppressor ............................................................................................ 40

5.6.2.3. SMAD4 ............................................................................................................................. 40

5.6.3. Genetically engineered mouse models (GEMMs) ............................................................... 41

5.6.4. Signalling pathways associated with PDAC ........................................................................ 42

5.6.5. Pancreatic Cancer Stem Cells ............................................................................................ 43

5.6.6. Novel mutations to define PDAC subgroups ....................................................................... 44

5.7. ATAXIA-TELANGIECTASIA MUTATED (ATM) ..................................................................... 47

5.7.1. ATM protein kinase structure: ............................................................................................ 48

5.7.2. Canonical ATM signalling ................................................................................................ 49

5.7.3. Non-canonical ATM signalling .......................................................................................... 50

5.7.4. Germline mutations and somatic mutations in ATM: ....................................................... 51

6. HYPOTHESIS AND AIMS ......................................................................................................... 52

6.1. THE ROLE OF Tbx3 AT THE TRANSITION FROM PLURIPOTENCY TO LINEAGE COMMITMENT .................................................. 52

6.2. A PROTOCOL TO GENERATE DEFINITIVE ENDODERM FROM HUMAN IPSCs .......................... 52

6.3. THE ROLE OF ATM IN PDAC USING A GEMM ...................................................................... 53

7. BIBLIOGRAPHY .......................................................................................................................... 54

8. APPENDIX .................................................................................................................................... 66

8.1. TBX3 DIRECTS CELL-FATE DECISIONS TOWARD MESENDODERM ...................................... 67

8.2. DEFINITIVE ENDODERM FORMATION FORM PLUCKED HUMAN HAIR-DERIVED INDUCED PLURIPOTENT STEM CELLS AND SK CHANNEL REGULATION .............................................................. 107

8.3. LOSS OF ATAXIA TELANGIECTASIA-MUTATED ACCELERATES PANCREATIC CANCER FORMATION AND EPITHELIAL-MESENCHYMAL TRANSITION .................................................. 122

9. STATUTORY DECLARATION ......................................................................................................... 149

10. ACKNOWLEDGEMENTS ............................................................................................................. 150

11. CURRICULUM VITAE .................................................................................................................. 151
2. Abbreviations

2i 2 inhibitors
ADM Acinar-to-ductal metaplasia
AP axis Anteroposterior axis
ATM Ataxia-telangiectasia mutated
AVE Anterior visceral endoderm
bHLH basic helix-loop-helix
BMP Bone morphogenetic protein
BRCA Breast cancer, early onset
c-Myc Myc proto-oncogene protein
CAC Centrocinar cell
CDKN2A Cyclin dependent kinase 2a / p16
CHASM Cancer-specific High-throughput Annotation of Somatic Mutations
CSCs Cancer stem cells
CT Computed tomography
Cxcr4 Chemokine receptor type 4
DE Definitive endoderm
DVE Distal visceral endoderm
E Embryonic day
EB Embryoid body
EGC Embryonic germ cell
EGF Epidermal Growth Factor
EMT Epithelial-to-mesenchymal transition
EpiSC Epiblast stem cells
ePTF1 Embryonic-type PTF1
ERK Extracellular signal-regulated kinases
ESC Embryonic stem cell
ExEn Extraembryonic
FGF2 Fibroblast growth factor
FPC Familial Pancreatic Cancer
FRT Flippase recognition target
FTDA FGF, TGF, Dorsmorphin, Activin A
G Grade
GDP Guanosine diphosphate
GEMM Genetically engineered mouse model
Gsc Goosecoid
GSK-3β Glycogen synthase kinase 3 beta
GTP Guanosine-5′-triphosphate
H-Ras Harvey-Ras
hESCs Human embryonic stem cells
ICM Inner cell mass
IDX1 Islet/duodenum homeobox 1
Indels Insertion/deletions
lpf1 Insulin promoter factor 1
3. Preface

This thesis is based on experimental work performed in the Department of Internal Medicine 1 Universitätsklinikum Ulm, Germany from April 2011 to June 2015. The project supervisor was Prof. Michael Kühl (Institute for Biochemistry and Molecular Biology, University Ulm, Germany) while most of the experiments were performed in the laboratory of PD. Dr. Alexander Kleger (Department of Internal Medicine 1, Universitätsklinikum Ulm, Germany).

The thesis is based on three publications:

**Paper I:**

“TBX3 Directs Cell-Fate Decision toward Mesendoderm”

* equal contribution.

**Paper II:**

“Definitive Endoderm Formation from Plucked Human Hair-Derived Induced Pluripotent Stem Cells and SK Channel Regulation.”
Paper III:

“Loss of Ataxia telangiectasia-mutated accelerates pancreatic cancer formation and epithelial-mesenchymal transition.”

Ronan Russell, Lukas Perkhofer, Stefan Liebau, Qiong Lin, André Lechel, Fenja M. Feld, Elisabeth Hessmann, Jochen Gaedcke, Melanie Güthle, Martin Zenke, Daniel Hartmann, Guido von Figura, Stephanie E. Weissinger, Karl-Lenhard Rudolph, Peter Möller, Jochen K. Lennerz, Thomas Seufferlein, Martin Wagner, Alexander Kleger.

4. Summary

Pluripotent stem cells are a powerful tool in which to dissect complex developmental questions. They provide a unique system, which can be easily manipulated in order to allow for detailed characterization of the pluripotent state, cell lineage commitment and organ development. Thus the role of individual or multiple genes and signalling pathways can be investigated in a controlled and time-sensitive manner to investigate their contribution to cell fate determination. The pancreas is a glandular organ derived from the definitive endoderm. It is composed of two functionally and morphologically distinct compartments and its primary functions are to regulate food digestion and maintain glucose homeostasis through either the exocrine or endocrine cell populations. There are a number of diseases associated with the pancreas including pancreatitis, genetically determined syndromes, diabetes and pancreatic cancer. In recent years, there has been a shift in our understanding of the complexity and heterogeneity of such diseases. It has been shown that several genes and signalling pathways implicated in normal pancreatic development processes including proliferation, morphogenesis and differentiation/maturation also have precise roles in maintaining normal homeostasis within the adult organ. Such knowledge has been garnered from embryological studies, genetically engineered mouse models (GEMM) and the use of embryonic stem cells (ESCs) or patient-derived induced pluripotent stem cells (iPSC). Thus, it appears that a very effective way to decipher the role of a particular gene or signalling pathway in a defined cell population is to use complementary approaches to delineate how it affects the normal development of that organ, its role in the normal adult organ and finally how it may contribute to disease pathophysiology.

The present work implements these different approaches to decipher the roles of different signalling pathways, genes and proteins in a context dependent manner, in order to better understand their roles in early cell lineage commitment, differentiation or disease progression respectively.

(1) In the first study, the role of the T-box factor Tbx3 in early cell lineage commitment of mouse ESCs was studied. We identified that TBX3 is dynamically expressed during the transition from pluripotency to the onset of specification of the mesendoderm lineages in differentiating embryonic stem cells (ESCs) in vitro and in developing mouse and
Xenopus embryos in vivo. Moreover, we show that Tbx3 directly activates key lineage specification factors such as Sox17 and Eomes, while concomitantly enhancing paracrine Nodal/Smad2 signalling. Taken together, we show that Tbx3 acts as an upstream regulator of early mesendoderm lineage commitment.

(2) In the second study, we utilized a refined protocol in order to generate iPSCs from an ectoderm derived cell source, namely, keratinocytes. Subsequently, we established highly efficient protocol for generating definitive endoderm from the newly derived iPSCs using a combination of growth factors and small chemical inhibitors. Finally we characterized the expression of small and intermediate conductance, calcium-activated potassium (SK) channels during the differentiation process.

(3) The final study investigates the role of Ataxia telangiectasia-mutated (ATM) in pancreatic cancer formation and progression using a genetically engineered mouse model. Our data reveals that loss of ATM promotes early neoplastic lesion formation and subsequently a faster disease progression. The phenotype associated with ATM-targeted mice includes a strong fibrotic reaction in the pancreas, altered BMP4 signalling coupled with enhanced epithelial-to-mesenchymal transition (EMT). We find a strong correlation with between our mouse model and a cohort of human PDACs, whereby low levels of ATM are associated with more aggressive disease subtypes and reduced survival.
5. Introduction

5.1. Pluripotency in mice and men

Pluripotency is a characteristic of embryonic stem cells (ESCs) that describes their ability to self-renew indefinitely, while maintaining the potential to differentiate into all cell types to make a whole organism, including germ cells \(^1\). Importantly, such cells alone cannot give rise to a full organism due to the absence of the extra-embryonic lineage, which arises from the trophectoderm. Therefore, they require a host blastocyst, which provides the necessary extra-embryonic structures required for subsequent development \(^2\). Pluripotent stem cells were first described in 1981 by two independent groups, who isolated cells form the inner cell mass (ICM) of embryonic day (E) 3.5 mouse blastocysts \(^3,4\). Such cells require special culture conditions in order to maintain them in the pluripotent state and the cell culture conditions have gradually been refined over recent years as described in more detail below. In addition, the cell state represented by these in vitro counterparts represents a very transient cell fate that occurs in vivo exclusively in the ICM of the developing blastocyst \(^5\). Therefore it is not surprising that subsequently, other promiscuous cell states have been identified, characterized and adapted to grow within cell culture in vitro \(^2\). These include epiblasts stem cells (EpiSCs) that represent cells obtained from the epiblast of E5.5 – E6.5 blastocysts \(^6,7\); primordial germ cells, which represent embryonic germ cell (EGCs) isolated at E10.5-11.5\(^8\); and gametes, which are spermatogonial germ stem cells (maGSCs) \(^9\). In addition, evidence now suggests that mouse ESCs (mESCs) can be derived from different stages of pre-implantation mouse embryos \(^10\).

Pluripotency can be assessed in several different ways \(^11\). Morphologically, mESCs grown on a feeder layer of inactivated mouse embryonic fibroblasts (MEFs) display tightly associated round clusters of cells and on a molecular level, show high levels of genes such as Oct3/4, Sox4 and Nanog as found in the E3.5 ICM of the developing mouse embryo. Alkaline phosphatase activity is another marker for pluripotent cells, as is Stage specific embryonic antigen-1 (SSEA-1). The genome of mESCs is hypo-methylated, in line with the methylation pattern of the genome of early blastocysts \(^12\). Interestingly, the promoter regions of the core pluripotency transcription factors (TFs) including Nanog, Klf, Sox2 and Oct3/4 are demethylated, allowing for robust expression.
in the pluripotent state. Upon differentiation, there is a rapid demethylation of genes associated with lineage commitment\textsuperscript{13,14}. More stringent tests for pluripotency include formation of teratomas that express all three germ layers\textsuperscript{15} or embryo aggregation assays in which mESCs are combined with healthy 2n or 4n blastocysts to give rise to chimeric mice\textsuperscript{16}. The ‘gold standard’ for testing pluripotency is injection of ESCs into the ICM of developing blastocysts to test for chimera formation\textsuperscript{17}. However this is a more expensive and time-consuming procedure and generally used only in particular circumstances where ESCs have been genetically manipulated or on newly established cell lines\textsuperscript{18}.

Similar techniques have been applied to human blastocysts, which has resulted in the isolation and establishment of human ESCs (hESCs)\textsuperscript{19,20}. These have slightly different characteristics to mESCs and require a different cocktail of growth factors and cytokines in order to maintain them \textit{in vitro}\textsuperscript{21}. Interestingly, hESCs are more similar in behaviour to mouse EpiSCs than mESCs. It is thought that mESCs represent the most naïve stage of pluripotency, possibly owing to their earlier stage of derivation, while mouse EpiSCs and hESCs have attributes that have resulted in the term “primed” pluripotent stem cells. Naïve pluripotency is sometimes referred to as the “ground state” of pluripotency and represents the developmental potential of the pre-implantation epiblast\textsuperscript{1}. This is highlighted by the ability of isolated single mouse epiblast cells microinjected into another blastocyst, to contribute to all cell lineages\textsuperscript{22}.

From E3.5 – 4.5 the blastocyst continues to grow and the ICM segregates into two separate layers, the primitive endoderm\textsuperscript{23} and the epiblast through a process of epithelialization. At this point, implantation occurs such that the blastocyst attaches to the uterine wall and undergoes dramatic changes in size and shape. The blastocyst becomes asymmetric, forming a cavity within the epiblast and extending along the proximal-distal axis to form the egg cylinder\textsuperscript{24}. The extraembryonic tissue settles as a cup-shaped layer of epithelial cells at the proximal pole. Moreover the distal visceral endoderm (DVE) moves to the anterior pole of the embryo to form the anterior visceral endoderm (AVE)\textsuperscript{25}. The epiblast is then subject to inductive factors emanating from the adjacent yolk sac and trophoblast tissue\textsuperscript{26}. Cell lines derived from the epiblast of these post-implantation embryos require different culture conditions to mESCs. Although there are some similarities, such as the expression of markers including Sox2, Nanog and
Oct3/4 and further and the ability to form teratomas when transplanted in vivo. However, there are number of important molecular and phenotypic differences, some of which are outlined in table 1. These include inactivation of one X chromosome, low clonogenicity, and inability or very poor ability to form chimeras \[^{27,28}\]. In addition, some genes are exclusively associated with the ground state of pluripotency such as Tbx3 and Klf4, and in contrast to the naïve state of mESCs, which utilize the distal enhancer to control Oct4, expression is driven by the proximal enhancer in EpiSCs \(^{13}\).

**Table 1. Comparison of Naïve and Primed Pluripotent States.** Reprinted from Naïve and Primed Pluripotent States, Nichols, J. and A. Smith, Cell Stem Cell, 2009. 4(6): p. 487-92, with permission from Elsevier\(^{1}\).

<table>
<thead>
<tr>
<th>Property</th>
<th>Ground State</th>
<th>Primed State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embryonic tissue</td>
<td>early epiblast</td>
<td>egg cylinder or embryonic disc</td>
</tr>
<tr>
<td>Culture stem cell</td>
<td>rodent ESCs</td>
<td>rodent EpiSCs; primate “ESCs”</td>
</tr>
<tr>
<td>Blastocyst chimaeras</td>
<td>yes</td>
<td>no(^a)</td>
</tr>
<tr>
<td>Teratomas</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Differentiation bias</td>
<td>none</td>
<td>variable</td>
</tr>
<tr>
<td>Pluripotency factors</td>
<td>Oct4, Nanog, Sox2, Klf2, Klf4</td>
<td>Oct4, Sox2, Nanog</td>
</tr>
<tr>
<td>Naive markers(^b)</td>
<td>Rex1, NrOb1, Fgf4</td>
<td>absent</td>
</tr>
<tr>
<td>Specification markers</td>
<td>absent</td>
<td>Fgf5, T</td>
</tr>
<tr>
<td>Response to Lif/Stat3</td>
<td>self-renewal</td>
<td>none</td>
</tr>
<tr>
<td>Response to Fgf/Erk</td>
<td>differentiation</td>
<td>self-renewal</td>
</tr>
<tr>
<td>Clonogenicity</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>XX status</td>
<td>XaXa</td>
<td>XaXi</td>
</tr>
<tr>
<td>Response to 2i</td>
<td>self-renewal</td>
<td>differentiation/death</td>
</tr>
</tbody>
</table>

\(^{a}\) Not applied to primate cells.  
\(^{b}\) Representative examples.

In recent years, there has been a large effort to reconcile the differences between the naïve and primed pluripotent states. This includes the generation of EpiSCs from ESCs
**in vitro** and vice versa. In terms of human ESCs, a number of groups have reported the conversion of human ESCs (hESCs) from the conventional primed state to the naïve state using different combinations of growth factors and small chemical inhibitors. However, the lack of corresponding host blastocysts to provide extraembryonic structures as well as ethical concerns means the true test of naïve pluripotency in hESCs is relatively limited compared to those of mESCs. In addition, most efforts to date have relied on re-setting primed hESCs and thus maybe considered a synthetic product of genetic intervention. Therefore, derivation of equivalent naïve cells from the human epiblast would provide a new benchmark within the field.

In 2006, Shinya Yamanaka first reported the successful reprogramming of mouse somatic cells into a pluripotent state. This was achieved by identification of four genes: Octamer binding transcription factor 3/4 (Oct3/4), SRY (sex determining region Y)-box 2 (Sox2), Krueppel-like factor 4 (Klf4) and c-Myc, which when virally transduced into mouse embryonic fibroblasts, promoted acquisition of the morphological and molecular features of mESCs. Such cells were termed induced pluripotent stem cells (iPSCs), and this technique has been shown to work in all somatic cell types tested to date. Further characterization of these cells demonstrated that they essentially behave in a parallel manner to ESCs and express high endogenous levels of the core pluripotency transcriptional network. Soon after this, similar results were obtained using human somatic cells. Since then, there has been a seismic increase in the amount of knowledge gained about the reprogramming process including at both the genetic and epigenetic levels. Importantly, all these studies were performed in vitro and only recently, Abad et al. demonstrated that reprogramming of somatic cells to iPSCs could be achieved in vivo. By generating reprogrammable mice carrying a doxycycline-inducible reprogramming cassette encoding the four reprogramming factors, the authors could show that adult somatic cells reprogrammed in vivo and acquired totipotent characteristics that were passed on by the zygote and its immediate daughter cells. iPSCs have enormous advantages over the use of hESCs, such as the availability of material, and the potential for drug screening for disease specific iPSCs. In particular, iPSCs are attractive in terms of moving towards the clinic, as they can be generated in a patient specific manner, differentiated in vitro a the required cell type and then re-implanted back to the same patient, thus reducing the risk of immune rejection.
compared to hESCs which are not histo-compatabile. These rapid advances in mouse and human ESC research have unearthed huge new potential and essentially diversified the fields of developmental biology, cancer research and regenerative medicine.

It was subsequently shown that the four Yamanaka factors may be substituted by other transcription factors. However, transduction of viral transgenes poses a risk of cancer formation and thus is not well suited for clinical applications. Therefore there has been a large focus on using alternative methods in order to reprogram somatic cells to a pluripotent state. Delivery of messenger RNAs (mRNA) and microRNAs, small molecules, piggy-back-transposons and episomal vectors has all shown promising results in various experimental set-ups. In addition, the direct trans-differentiation of one cell type to another has been applied, although the intermediate steps by which cells lose characteristics of one cell type and gain properties of another are less well defined than starting from a pluripotent cell population. Moreover, other methods have since been exploited in order to generate pluripotent cells from somatic cell types. Somatic cell nuclear transfer (SCNT) involves the injection of the nucleus from a somatic cell into an enucleated oocyte. The somatic cell nucleus undergoes reprogramming in the host egg cell and development will progress a normal, generating a blastocyst. Another approach is the fusion of ESCs and somatic cells, which results in hybrid cells that maintain a stable tetraploid DNA content and behave with characteristics of ESCs. In most cases, such approaches are developed using the murine system and become gradually refined to work in the human setting.
5.1.1. Signalling pathways associated with pluripotency

The *in vitro* maintenance of mESCs and hESCs in a pluripotent state relies on different requirements. Several signalling axes work synergistically in order to prevent differentiation and maintain pluripotency. Traditionally, mESCs were co-cultured in serum containing medium with mouse embryonic fibroblasts (MEFs). It is now known that these “feeder” cells secrete leukaemia inhibitory factor (LIF) \(^{51}\), which acts upstream of the TF - signal transducer and activator of transcription 3 (Stat3), which promotes ESC viability and prevents differentiation. Further biochemical and genetic screening studies identified that activation of the ERK pathway through various intrinsic and extrinsic stimuli promoted differentiation of mESCs. Interestingly, the common dependence of mESCs and diapause embryos on the cytokine LIF has been interpreted by some to indicate that murine ESCs employ a diapause-like program for their maintenance *in vitro* \(^{5,52}\). In traditional mESC cultures in LIF/serum/MEFs, other signalling pathways are implicated. In addition to its role in regulating Stat3 signalling \(^{53}\), LIF also acts to inhibit PI3K and induce ERK phosphorylation \(^{54,55}\). Moreover, bone morphogenetic protein 4 (BMP4) is sufficient to replace the requirement of serum in order to maintain ESC pluripotency. This occurs in combination with LIF and it has been shown that BMP4 acts to inhibit the mitogen-activated protein kinase (MAPK) pathway to prevent extracellular signal-regulated kinase (ERK) phosphorylation \(^{56}\). This led to the discovery by Austin Smith and colleagues that chemical inhibition of the MEK/ERK and glycogen synthase kinase 3 β (GSK-3β) signalling pathways promoted a robust cell culture system for the maintenance of mESCs in the ground state \(^{57}\). The combined used of these two inhibitors was termed ‘2i’ and comprises PD0325901, which selectively inhibits MAPK/ERK mediated Klf2 phosphorylation and Chir99021 which selectively inhibits GSK-3β, enabling transcriptional activation of pluripotency-associated genes \(^{57-59}\). Importantly, addition of 2i during the isolation and cultivation of mESCs from refractory strains significantly improves mESC derivation efficiency and has also been reported to improve iPSC generation \(^{27,60,61}\). Thus the use of the chemical combination 2i is preferable due to cost effectiveness and less batch-to-batch variation compared to the use of animal derived products such as serum.

In a similar manner, hESCs are also often cultured on top of a MEF feeder layer or in MEF conditioned medium. However, while such conditions are fine for maintenance of
pluripotency, it somewhat complicates the directed differentiation of cells to a particular cell lineage due to the presence of contaminating MEFs and the fact that such cells grow in tight clusters. Therefore, different efforts have highlighted alternative ways in which to culture hESCs in feeder-free conditions and whereby they can be easily dissociated into single cells to allow for more reproducible starting populations for downstream experiments. Much of this knowledge was obtained through identification of factors present in MEF conditioned medium. However, many of these commercially available media are expensive and their compositions are either not fully disclosed or are relatively complex. Instead, one of the most basic media formulations which supports hESCs maintenance is termed FTDA medium. With the aim of only including components that are necessary to promote growth and self-renewal of hPSCs, the group of Boris Gerber identified a set of growth factors and small chemical inhibitors that achieves this and presents a cost-effective alternative for expansion of hPSCs. FTDA is an acronym for fibroblast growth factor 2 (FGF2), transforming growth factor β1 (TGFβ1), Dorsomorphin (DM) and Activin A. FTDA represent factors identified in MEF conditioned medium, which are members of the TGFβ family of ligands and were proposed to promote self-renewal in hESCS. DM is a small molecule inhibitor of the BMP pathway and it was effectively shown that by adding low amounts of DM to defined hESC media, spontaneous extraembryonic differentiation could be abolished. Therefore this defined media effectively results in cooperative support of hESC growth and self-renewal.

Taken together, these studies have rapidly advanced the field of stem cell research. By taking advantage of knowledge previously existing in the field of developmental biology and by modulating and titrating different signalling pathways, it is now possible to isolate, maintain and expand pluripotent stem cells from mice and men in relatively defined conditions.
5.1.2. Pluripotency associated transcription factor network

There are a number of criteria by which pluripotency can be assessed as outlined above. These include signalling molecules, microRNAs and epigenetic modifiers. Interestingly, one of the most well studied aspects of pluripotency studied to date is the contribution of different TFs, which integrate to safeguard the pluripotent state and appear to be on top of the hierarchy of pluripotency associated regulatory controls. This has resulted in the term “pluripotency-associated transcription factor network” which represents a set of core TFs whose expression has been shown to be necessary in order for cells to exist in a pluripotent state. These include Oct4, Nanog and Sox2, and a complex system of coregulatory and autoregulatory mechanisms have been shown to link these three TFs to establish a self-reinforcing circuit. Individually, Oct3/4, Nanog and Sox2 have been shown to be critical for establishment of pluripotent epiblast cells in the blastocyst. Loss of any individual factor results in spontaneous differentiation of ESCs. Thus it is widely believed that each TF acts to suppress functional expression and activity of lineage specification factors such as Cdx2, Gata4, Gata6, and that loss of any one individual factor shifts the balance away from a state of pluripotent homeostasis and towards lineage commitment.

Oct3/4 (Aliases: Otf3, Otf4, POU5F1) is a mammalian homeodomain transcription factor and member of the POU (Pit-Oct-Unc) family. It has effectively been shown that Oct3/4 is indispensable for induction and maintenance of pluripotency as Oct3/4-deficient mouse embryos fail to form a pluripotent ICM, resulting in embryonic lethality due to uncontrolled differentiation of the extraembryonic trophoblast lineage. Interestingly, Oct4 interacts with several signalling pathways associated with pluripotency, including LIF/Stat3. Moreover, together with Sox2, Oct4 binds a unique Oct/Sox motif on the promoter region of many other genes associated with pluripotency.

Sox2 is a member of the HMG (high-mobility group) box proteins. Sox2 is mainly expressed in the ICM, epiblast and TE of pre-gastrulation embryos. Sox2 homozygous null embryos are lethal soon after implantation and as the ExEn develops normally in Sox2 null chimeras rescued with wild-type ES cells, this suggests that this is the only lineage with no cell-autonomous requirement for Sox2.
Nanog is a homeodomain TF and interestingly, its overexpression is sufficient to maintain pluripotency of mESCs independently from LIF/STAT3, albeit at a reduced self-renewal capacity level. Notably, Nanog influences the binding of Oct3/4 and Sox2 at the Oct/Sox motif. Nanog is first expressed at the morula stage and becomes restricted to the ICM and finally subsequently the epiblast. Nanog null mice failed to generate an epiblast and only produced parietal endoderm-like cells. Interestingly, some recent studies have shown that Nanog is dispensable for acquisition of pluripotency during the reprogramming process.

In addition to this trinity of core factors, many other TFs have been implicated with maintenance of pluripotency. Interestingly, many of these have been sub-classified by their expression in naïve ESCs as opposed to the more primed cells and they are often integrated into the core pluripotency network via direct and/or indirect connections as outlined in table 2 below.
Table 2. Connections between ancillary and core pluripotency transcription factors. Modified from [80].

<table>
<thead>
<tr>
<th>Gene</th>
<th>Transcriptional interaction with core factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esrrb</td>
<td>Interacts with Oct4, Nanog, Sox2 and Ncoa3; Activate Nanog, Oct4 and other genes; Target of Nanog and Tcf3</td>
</tr>
<tr>
<td>Nr5a2</td>
<td>Interacts with Dax1 and Rary; Activates Oct4 and Nanog; Direct target of Oct4, β-Catenin and Tcf3.</td>
</tr>
<tr>
<td>Dax1</td>
<td>Interacts with Nanog, Oct4 and Nr5a2; Activate Oct4 expression; Target of Stat3, Oct4, Esrrb, Sox2, Nr5a2, Nanog and β-catenin.</td>
</tr>
<tr>
<td>GCNF</td>
<td>Repress Oct4 and Nanog upon differentiation.</td>
</tr>
<tr>
<td>Klf4</td>
<td>Interacts with Oct4 and Sox2; Activate Nanog, Esrrb and other genes; Downstream target of LIF/Stat3 signalling.</td>
</tr>
<tr>
<td>Sall4</td>
<td>Interact with Nanog, Oct4, Sox2, Esrrb, Dax1, MTA2 and Nac1; Activate Oct4 expression; Linked to TGF-β and WNT signalling through Usp9X and Cxxc5; Involved in transcriptional repression, cell cycle regulation(via binding to cyclin D1).</td>
</tr>
<tr>
<td>FoxD3</td>
<td>Activate Nanog through cooperating with Oct4; Target of Oct4; High level of FoxD3 can inhibit itself.</td>
</tr>
<tr>
<td>FoxO1</td>
<td>Essential for human ESC maintenance; Activate Oct4 and Sox2 in human ESCs.</td>
</tr>
<tr>
<td>Zfp206</td>
<td>Interact with Oct4, Sox2, Zscan4 and Zfp110; Activate Nanog and Oct4; Target of Oct4 and Nanog.</td>
</tr>
<tr>
<td>Zfp296</td>
<td>Activate Oct4; Enhance reprogramming.</td>
</tr>
<tr>
<td>Zfp281</td>
<td>Interact with Oct4, Sox2 and Nanog; Activate Nanog.</td>
</tr>
<tr>
<td>Rex1</td>
<td>Target of Oct4, Sox2, Nanog, Klf4, Dax1 and Nacl; Maintains X-activation, imprinting, cell cycle, and mitochondrial fission in ESCs.</td>
</tr>
<tr>
<td>Zscan4</td>
<td>Overexpression or knockdown Zscan4 did not alter the transcriptome dramatically; Transient expression of Zscan4 lead to telomere elongation and can restore the developmental potency of ESC.</td>
</tr>
<tr>
<td>Dapp2 &amp;</td>
<td>Essential for ESC maintenance;</td>
</tr>
<tr>
<td>Tbx3</td>
<td>Activate Nanog, Oct4, Sox2, Sall4, Lefty1, Lefty2, Zfp42, Klf2, Klf4, Klf5, n-Myc and c-Myc; Target of Nanog and Tcf3; Partially activated by PI3-kinase but inhibited by MAP-kinase; Repress the expression of cell cycle regulators NFκBIB and p14ARF.</td>
</tr>
<tr>
<td>PRDM14</td>
<td>Activate Oct4 in human ESCs; Repress differentiation-related genes; Interact with PRC2 complex to repress gene expression.</td>
</tr>
<tr>
<td>L1td1</td>
<td>Interacts with Lin28 to modulate levels of Oct4; Target of Oct4, Sox2, and Nanog in human ESCs.</td>
</tr>
<tr>
<td>Utf1</td>
<td>Direct target of Oct4 and Sox2; Involved in chromatin-associated transcriptional repression.</td>
</tr>
</tbody>
</table>
5.2. T-box factors

The T-box family comprises a set of evolutionary important genes defined by a common DNA binding motif, known as the T-box \(^ {81}\). To date, 19 individual T-box genes have been identified and these transcription factors have been classified into 5 subfamilies based on phylogenetic analysis \(^ {82}\). Much of the current knowledge of these transcription factors has been derived from mutations discovered within these genes, the first being Brachyury (T), which was identified to cause truncated tails in mice \(^ {83}\). Since then, loss of function studies in mice and other organisms have established that this family of genes is important particularly for normal embryonic development and later during organogenesis \(^ {82}\). In line with this, mutations within different family members are associated with a wide range of developmental abnormalities and syndromes \(^ {84}\).

The seminal work on T-box factors identified the T-box DNA binding element (TBE) or sequence within the T gene and consists of a palindromic sequence (5'-AGGTGTGAAATT-3') that binds to specific sequences in the promoter of particular target genes. T binds this sequence as a dimer and each monomer of T binds half of \(^ {85}\) the sequence, termed the T-half site. In recent years, it has been established that all T-box family members are capable of binding the T-half site as monomers, with different preferences for T-half sites, which diverge in terms of orientation, numbers, and spacing \(^ {86,87}\). Closely related T-box factors can be grouped based on phylogenetic analysis as outlined in figure 1 below. It is believed that such variations within the T-half site ensure a greater specificity in order to bind the promoter of down-stream target genes. Notably, the T-box proteins bind within the minor and major grooves of the T-domain structure. Moreover, T-box family members comprise different activator and repressor domains, generally located at the C-terminus, and the use of such domains is thought to be promoter context-dependent, as for Tbx2 or Tbx3 which comprise both activator and repressor domains \(^ {85,88}\).
5.2.1. T-box factor 3

Tbx3 belongs to the Tbx2 subfamily. Three different transcript variants encode different protein isoforms, which exist due to alternative splicing of the Tbx3 gene. While the full structure of one variant has not yet been determined, the Tbx3 and Tbx3+2a isoforms have been reported to have alternative binding preferences. In humans, Tbx3 plays a critical role in normal development. Heterozygous mutations within Tbx3 including point mutations or insertion/deletions (indels) result in ulnar-mammary syndrome (UMS), a rare autosomal dominant condition that exhibits incomplete penetrance. There is a high degree of variability within the disease phenotype, but commonly comprises of congenital defects of the upper extremities of the body, including the limbs and hypoplasia of the apocrine and mammary glands. Moreover, there are genital abnormalities such as reduced ability to lactate and uterine defects in females and delayed puberty and genital hypoplasia are commonly found in males. Other abnormalities associated with this disease include defects of the heart, endocrine system and teeth. Tbx3 has also been reported to be mis-expressed in a variety of
cancers. Interestingly, various mouse models in which Tbx3 has been targeted have identified various and subtle differences due to complete loss of Tbx3 or the presence of a truncated protein. This is in line with the vast array of phenotypes observed in patients with mutations within Tbx3 and thus it is thought that in a similar manner to other T-box proteins, Tbx3 activity is highly dosage sensitive. Homozygous loss of Tbx3 is embryonic lethal and there is a broad time-window during which Tbx3 null mice die (E9.5 – E16.5). This is primarily due to major deficiencies in yolk sac and limb development as well as a smaller sinoatrial node, atrioventricular blockage, blockage and ventricular bradycardia.

Recent work has identified T-Box 3 gene (Tbx3, or D5Ertd189e) to be the only T-box family member expressed in the morula as well as in the inner cell mass of preimplantation embryos and in turn in pluripotent ESCs. Interestingly, Tbx3 has also been shown to have a role in regulation of Zscan4+/2 cell embryo state. It has been reported that removal of Tbx3 from mESCs results in loss of pluripotency and rapid onset of differentiation. Therefore, Tbx3 is a unique member of the T-fox family due to its role in maintenance of pluripotency and interestingly it is considered to be a marker for ground state pluripotent mESCs. Other evidence for a direct role of Tbx3 in pluripotency comes from studies of iPSCs, whereby Tbx3 significantly improves iPSC quality shown by high efficiency germ-line transmission. Moreover, detailed characterization of mESCs have revealed that Tbx3 regulates key members of the pluripotency associated transcription factor network such as direct binding of the Nanog and Oct3/4 promoters, and mediates LIF/STAT signalling. As mentioned above, Tbx3 was recently classified as a gene that promotes mesendodermal derivatives. Moreover, Tbx3 was identified to be enriched within a population of self-renewing endodermal progenitor cells from hESCs and is specifically expressed in multipotent hepatic progenitor cells, ‘hepatoblasts’, within the developing mouse liver. Finally, Tbx3 was recently shown to associate with the histone demethylase Jmjd3 at the enhancer element of the Eomes locus, thus allowing for enhancer-promoter interactions, and activation of a core transcriptional network required for endodermal differentiation. Taken together, these studies suggest that Tbx3 is required for maintenance of pluripotency and also functions at different stages during the transition from pluripotency to cell lineage commitment.
5.3. Pluripotency and cell lineage commitment

Embryonic development can be classified into 4 stages: (1) Cleavage, (2) Patterning, (3) Differentiation and (4) Growth. Early embryonic development in Xenopus, zebrafish and mice has been well documented and the emergence of the different germ layers: endoderm, mesoderm and ectoderm, has been extensively studied in terms of timing and the major underlying molecular mechanisms and signalling cues. Until the wide use of hESCs in many labs across the world, most of our knowledge of early embryonic development was derived from animal models, in particular the mouse. However, in the last decades and with the advent of hESCs and the iPSC technology, there has been a rapid shift to translate these previous findings and knowledge to humans and the same principles have been widely adapted to mimic differentiation of pluripotent stem cells in vitro.

5.3.1. Gastrulation and Endoderm specification

The endoderm is classically defined as the inner germ layer of the developing embryo, which gives rise primarily to the epithelium of the digestive tract and organs within close proximity. The endoderm germ layer is first specified during gastrulation in the mouse embryo. Gastrulation is a well-conserved evolutionary morphogenetic process occurring around day E6.5, whereby the embryo is characterised by regional differences in gene expression and signalling molecules. A Primitive Streak (PS) forms in a specific region of the Epiblast along the posterior axis of the embryo and signalling molecules including WNTs, TGF-β, Nodal, BMPs, and FGFs act in concert to allow for correct segregation of the germ layers. This primarily occurs through epiblast stem cells undergoing an epithelial-to-mesenchymal transition (EMT). The PS elongates and at the anterior tip, the node forms. The Node, together with the anterior visceral endoderm (AVE) regulates patterning of the embryo. Epiblast cells move through the PS and spread forward and laterally between the ectoderm and the AVE as outlined in figure 2.

In most species studied to date, with only a few exceptions, there is a separation of cells during gastrulation into progenitors which are poised to form ectoderm and a second bipotent progenitor cell type which specifies cells destined to make mesoderm and endodermal derivatives. This bipotent progenitor cell type is termed a mesendodermal progenitor population and it appears that some common signalling cascades induce both mesoderm and endoderm at the early stages of cell lineage commitment such as
Nodal/Smad2 signalling \(^{111,112}\). Upstream of Nodal, is the Wnt pathway and it has been demonstrated via genetic ablation studies in mouse that loss of Nodal or β-catenin results in primitive streak failure \(^{113,114}\). Later, the definitive endoderm replaces the visceral endoderm. The anterior epiblast generates the neuroectoderm and the ectoderm that covers the surface of the embryo \(^{25}\). In addition to Nodal and Wnt, other signalling molecules and pathways have been identified to act in a spatial and temporal manner to allow for correct segregation of the developing organs \(^{25}\). This highly organized and complex series of events ensure correct segregation of the germ layers and allow for subsequent organ specification and maturation \(^{24}\).

**Figure 2. Origin of stem cells in the mammalian embryo.** In this figure, the pluripotent cells of the embryo are tracked in green. From left to right, the embryo goes through a series of molecular and structural changes from the morula stage (E2.5) which holds a core of cells poised to generate the ICM at the E3-E4 blastula stage. After further segregation events, at E6.0, the embryo will begin the process of gastrulation that involves the specification of the three germ layers endoderm, mesoderm and ectoderm. In addition, the primordial germ cells (PGCs) are also specified during this stage. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology \(^{24}\), copyright (2005).
5.3.2. Spontaneous Differentiation

Given the developmental potential of pluripotent stem cells, an important advancement in our understanding of germ layer specification in mice and men, has been the use of both spontaneous differentiation and directed differentiation of pluripotent stem cells in vitro. In vivo, gastrulation marks the beginning of germ layer specification while this tightly regulated process ensures correct segregation of the 3 layers within a short-timeframe\textsuperscript{25}. In vitro, spontaneous differentiation to the 3 germ layers may be achieved by removal of one or more of the pluripotency associated signalling factors such as LIF (mESCs) or TGFβ (hESCs) or by the addition of compounds that promote differentiation, such as retinoic acid. Moreover, embryoid bodies (EBs) represent three-dimensional aggregates of pluripotent stem cells that undergo spontaneous differentiation towards the three germ layers and are an effective in vitro model in which to study early embryogenesis and cell fate decisions\textsuperscript{115-117}. However as the goal of many researchers is to generate pure populations of differentiated cell types, a major field of study is the directed differentiation of pluripotent stem cells from mice or men to a particular cell lineage such as pancreatic beta cells. For example, as the pancreas is an organ derived from the endoderm, it is therefore critical that pluripotent cells are initially directed to efficiently form endoderm at the cost of mesoderm and ectoderm\textsuperscript{118}.

5.3.3. Directed Differentiation

In vitro, although mESCs and hESCs can spontaneously differentiate into all three germ layers, in conventional cell cultures, endoderm is formed at very low efficiency\textsuperscript{119}. This is likely due to factors present in serum that promote ectoderm differentiation\textsuperscript{119}. Therefore recent efforts have aimed to generate relatively pure endoderm at high efficiencies, as prior efforts at generating endodermal derivatives such as pancreatic beta cells or hepatocytes were poorly characterized and suffered from lack of reproducibility, low efficiency and/or poor functionality\textsuperscript{120,121}. Thus an important aspect of differentiating pluripotent stem cells to mature adult cell types is now considered to be an in-depth characterization of the intermediate steps throughout the differentiation process\textsuperscript{118}. In order to achieve this, most strategies rely primarily on knowledge from animal models such as signalling cues and molecular makers in order to correlate the in vitro differentiated cells to counterparts found in vivo at defined time-points\textsuperscript{118}. 
Therefore most in vitro differentiation protocols aim to mimic the in vivo niche and direct cell lineage specification specifically to one of the three germ layers. The knowledge gained from such in vivo developmental studies have provided a blue-print which researchers have exploited in order to direct differentiation of pluripotent stem cells in vitro. Earlier studies on mESCs have identified some of the major factors required for the initialisation of mesendoderm in pluripotent stem cells and have paved the way for subsequent work in hESCs and iPSCs. Fluorescent reporter mESCs have also been advantageous in allowing for tracing of emerging distinct cell populations that arise early within the primitive streak 122-124. The laboratories of Doug Melton and Gordon Keller have been pioneering within this field and have led to the characterisation of an in vitro cell state reflective of mesendoderm, whereby expression of Eomesodermin (Eomes), Foxa2, Goosecoid (Gsc) and Brachyury (T) are upregulated and there is a reduction in the levels of the pluripotency associated transcription factors Oct3/4, Sox2 and Nanog 125,126. Interestingly, members of the core pluripotency TF network were recently reported to adopt new roles during the transition from maintenance of pluripotency to the onset of germ layer specification and cell lineage commitment. Thomson et al., described a “mesendoderm class embryonic stem cell gene” set which identified pluripotency TFs such as Nanog, Oct4, KLF5 and Tbx3 as having specific signatures associated with mesendodermal cell lineage specification, while Sox2 and RBPJ genes were associated with neuroectoderm lineage commitment 127. This work suggests an intimate link between maintenance of pluripotency and subsequent lineage commitment.

After mesendoderm specification, subsequent separation into both mesoderm and endoderm occurs. The expression of the transcription factors Sox17 and Foxa2 and the cell surface marker Cxcr4 can be used to identify endoderm, while mesoderm expresses both T and Mixl1 124,128. Various protocols have been meanwhile established in order to direct cell differentiation towards either of these populations. In addition, there is still a strong effort within the field to refine existing protocols in such a way as to make them cost effective, efficient and highly reproducible. Due to the high cost and limited availability biologically active Nodal protein 125, many protocols have made use of Activin as a surrogate for Nodal and the work of D’Armour et al., provided one of the most reproducible protocols for differentiating hESCs to definitive endoderm (DE) at high efficiency and subsequently, into pancreatic progenitor cells and pancreatic beta cells.
Briefly, endodermal differentiation protocols rely on Tgfβ signalling activation via Activin A and activation of Wnt signalling as shown in figure 3. This effectively patterns the cells to generate DE in a relatively efficient manner. Notably, there are many variations of this general protocol being used, and as our knowledge and understanding of hESC differentiation increases, they are constantly being refined. This includes titrating the different amounts of Nodal and Wnt signalling as well as the use of small chemical compounds to avoid the use of animal derived products.

Figure 3. Schematic of the directed differentiation procedure to generate pancreatic endocrine cells from hESCs through a series of defined intermediate stages. The differentiation protocol is divided into five stages and the growth factors, medium and range of duration for each stage are shown. This protocol orchestrates differentiation through five identifiable endodermal intermediates en route to production of hormone-expressing endocrine cells. Several corresponding markers characteristic of each cell population are listed in the bottom panel. Reprinted by permission form Macmillan Publishers Ltd (Nature Biotechnology) 118, copyright (2006).
5.4. Pancreatic development and regulating key factors

At the end of gastrulation, the definitive endoderm forms the epithelium of the primitive gut tube (PGT)\(^{25}\). The anterior gut tube gives rise to the foregut, while the posterior region gives rise to the hindgut\(^ {109}\). The initial morphological appearance of the pancreas occurs at day E9.5 in the mouse pancreas and approximately day 26 in human gestation, when there is a thickening of the dorsal side of the foregut epithelium. Concomitantly, two ventral buds emerge laterally from the epithelium, with one regressing before gut rotation. Afterwards, the dorsal and ventral buds fuse and the pancreatic epithelium branches in close proximity to the surrounding mesenchyme\(^ {109}\).

At this stage, around E12.5, multipotent progenitor cells have the ability to form all cell types of the pancreas\(^ {130}\). These highly proliferative cells migrate to the epithelial branches and differentiate into either pre-acinar cells at the tips, or bi-potent islet/ductal cells along the epithelial branches. Further lineage specification and maturation occurs to generate the complete adult pancreas between E15.5 and the time of birth at approximately E21\(^ {131}\).

The emerging pancreatic endoderm is initially marked by the expression of two transcription factors which are essential for normal pancreatic development and maturation: (1) the pancreas and duodenal homeobox gene-1 (Pdx1) and (2) the pancreatic transcription factor-1 (Ptf1a/p48)\(^ {130}\). Both of these factors have clearly defined roles in promoting lineage commitment of the endocrine and exocrine compartments, as shown by lineage tracing studies in mice\(^ {132}\)\(^ {133}\). However, during further specification both factors mark distinct lineages, namely the exocrine and endocrine pancreas as outlined in figure 4. The pancreatic and duodenal homeobox 1 gene (Pdx1) is first expressed around E8.5 in murine embryogenesis and is also known as insulin promoter factor 1 (Ipf1) or islet/duodenum homeobox 1 (IDX1)\(^ {134}\). Pdx1 marks a region of endoderm with intrinsic foregut competence and this area is further specified by precise intercellular signalling further subdivided by specific intercellular signalling proceedings that drive cells towards liver or pancreas specification\(^ {135}\). At E10, Pdx1 is uniformly expressed in the dorsal and pancreatic buds, but immediately after, its expression is down-regulated, with high levels re-emerging only in the beta cells from E11 onwards. Notably in the adult pancreas, Pdx1 is most abundantly expressed in the endocrine compartment particularly in beta cells, with little or no expression observed in
the exocrine compartment\textsuperscript{133}. Inactivation of this gene causes pancreatic organogenesis arrest in murine embryos and diabetes in adult mice as shown by conditional knockout studies\textsuperscript{136}. Heterozygous Pdx1 mice display an age-dependent worsening of glucose intolerance as well as an impaired glucose-stimulated insulin release and a similar role was demonstrated in humans, where haploinsufficiency leads to maturity onset of the young (MODY4), an early onset form of diabetes\textsuperscript{137}. Taken together, these studies highlight the crucial role of Pdx1 in pancreatic cell lineage specification. However, Pdx1 alone is not sufficient for orchestrating all events in normal pancreatic development and thus is thought to act in concert with other important factors, as early events of pancreatic morphogenesis take place even in the absence of functional Pdx1\textsuperscript{134}.

PTF1a (p48) is the α-subunit of the transcription complex pancreas-specific transcription factor (PTF1). Ptf1a is a basic helix-loop-helix (bHLH) protein that is highly expressed in the exocrine compartment of adult and human pancreas and is one of the primary transcription factors regulating acinar cell differentiation and maturation\textsuperscript{138}. In particular, the Ptf1a gene trans-activates 27 of the 28 key acinar secretory proteins, as well as genes implicated in nitrogen and energy metabolism, creatine biosynthesis and other constituents of the intracellular protein transport apparatus. The knockout of Ptf1a results in complete absence of exocrine pancreatic tissue, with some alteration found in the endocrine compartment\textsuperscript{138}. Moreover, Ptf1a also has a role in the brain, as knockout mice exhibit postnatal lethality due to central nervous system problems\textsuperscript{139}. In the developing mouse embryo, Ptf1a is activated shortly after Pdx1, appearing at 9.5 dpc in mice development and is required for the evagination of the ventral bud and the growth of the dorsal bud. Interestingly, the group of Chris Wright have effectively shown that expressing Pdx1 under the control of an extended Ptf1a promoter could restore pancreas formation to Pdx1 knockout mice\textsuperscript{132}. These studies suggest that Ptf1a expression in the early pancreatic buds is independent of Pdx1 function. The PTF1a acts in a trimeric complex comprising Ptf1a, a bHLH cofactor that binds DNA such as E12 and another subunit that binds TC-rich DNA sequences\textsuperscript{140}. In early embryogenesis, this subunit is Rbpj / Rbpjk /CSL /Cbf1 and the complete complex is designated embryonic-type PTF1 (ePTF1). During the onset of acinar differentiation, this complex undergoes a conformational change, whereby Rbpjl, a closely related parologue
replaces Rbpj and drives acinar cell maturation. Similar processes are thought to occur in humans as highlighted from limited studies of *ex vivo* culture of human fetal pancreas and the increasing use of pluripotent stem cells. 

Figure 4. Pancreatic lineage specification. Pancreas specification is first marked by the expression of the transcription factor Pdx1 within the foregut endoderm. As organogenesis proceeds, multipotent progenitor cells are marked by a combination of transcription factors (green). These progenitors segregate to form bi-potent trunk cells (red) that give rise to endocrine (purple) or exocrine (blue and orange) cell populations that are further defined by the expression of sub-sets of transcription factors.
5.5. Pancreas anatomy and physiology

The pancreas is a glandular organ composed of functionally and morphologically distinct compartments as shown in figure 5. The origin of the organ name derives from ancient Greek studies of human anatomy, whereby the lack of bone or cartilage led to the term ‘pan’ meaning ‘all’ and ‘creas’ meaning ‘flesh’. Located in the upper abdomen region, the human pancreas forms a well-defined organ typically weighing about 100 grams and 15-25 centimetres in length. However in lower organisms, the structure is less defined. The pancreas is connected to the duodenum via the ampulla of Vater whereby the primary pancreatic duct connects to the common bile duct. The primary function of the pancreas is to regulate food digestion and maintain glucose homeostasis through either the exocrine or endocrine cell populations.

Figure 5. Anatomy of the pancreas. The pancreas is comprised of separate functional units that regulate two major physiological processes: digestion and glucose metabolism. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Cancer copyright 2002.
5.5.1. The exocrine pancreas

The exocrine pancreas is composed of enzyme secreting acinar cells and ductal cells which transport the enzymes to the duodenum\textsuperscript{144}. Together, the exocrine portion of the pancreas comprises over 90\% of the whole organ. Terminally differentiated acinar cells are organized into functional cell clusters termed ‘Acini’ and are characterized by their pyramidal shape with a basal nucleus, a highly active, prominent Golgi complex, rough endoplasmic reticulum and numerous dark zymogen granules containing the enzymes located at the apical pole. Acini are located at the end of a mature ductal network comprised of four sub-compartments (1) the primary duct connecting to the duodenum; (2) interlobular ducts which connect to the primary duct; (3) intralobular ducts which drain the exocrine enzymes into the interlobular ducts; (4) intercalated terminal ducts also termed centroacinar cell (CACs) which are situated between the acini and the intralobular ducts.\textsuperscript{145,146}

Pancreatic acini are highly specialized structures responsible for synthesis, storage and regulated secretion of 28 individual digestive enzymes. Many such enzymes are released in a “pancreatic juice” of inactive precursors to the pancreatic ducts, from where they are then transported to the duodenum and become activated. The secretion process requires signals from hormones including secretin and gastrin but also occurs via neural stimuli, highlighting the complex regulation of this specialized cellular compartment\textsuperscript{147}.

The branched ductal network directs the pancreatic juices into the gastrointestinal tract and also produces bicarbonate that neutralizes stomach acidity. The CACs or intercalated duct cells reside adjacent to the acini and form a simple cuboidal epithelium, rich in mitochondria. As the duct extends, the CACs converge to form the intralobular ducts that serve a group of acini or lobes of the pancreas. Interlobular ducts become progressively larger, transitioning from a thin squamous epithelial layer, to cuboidal layers with intertwined goblet cells. These ducts connect the different lobes of the pancreas and merge into the primary duct. The main pancreatic duct contains columnar epithelium surrounded by rich connective tissue and merges with the common bile duct\textsuperscript{148}.
5.5.2. The endocrine pancreas

The endocrine compartment is composed of pancreatic islets of Langerhans that are responsible for glucose homeostasis and regulation of nutrient metabolism. The islets of Langerhans are polyclonal, compact cell clusters embedded among the exocrine tissue and consist of five individual hormone-secreting endocrine cell types: (1) the beta cells which secrete both insulin and the insulin antagonist - amylin, (2) glucagon-secreting alpha cells, (3) somatostatin-releasing delta cells, (4) ghrelin-producing e cells and (5) the pancreatic polypeptide-secreting (PP) cells \(^{144}\). These islets are located in close proximity to neurons, a mesoderm – derived stromal component and importantly, to blood vessels. Such proximity to the vasculature is essential in regulating hormone release in order to maintain glucose homeostasis within the body \(^{149}\).

5.6. Pancreatic ductal adenocarcinoma (PDAC)

Pancreatic cancer represents one of the most lethal types of cancer with a 5 year survival rate between 4 and 6% and accounting for more than 220,000 annual deaths worldwide \(^{150}\). Notably, Pancreatic Ductal Adenocarcinoma (PDAC) is the most common type of pancreatic neoplasm, accounting for more than 85% of cases \(^{151}\). Such tumours are classified due to their histology, which resemble pancreatic duct cells. The lack of early signs or symptoms of pancreatic cancer within patients are insidious and unspecific. In general terms, patients experience weight loss and abdominal and/or back pain. Obstruction of the bile duct can give rise to jaundice and in some cases, patients experience glucose intolerance. Less commonly, enlargement of the gall bladder may occur known as Courvoisier’s sign or Trousseau’s syndrome whereby migratory thrombophlebitis or venous thrombosis occur. This often results in delayed detection and subsequently, many patients being diagnosed at an advanced stage of disease \(^{150}\). Upon clinical suspicion of PDAC, computed tomography (CT) \(^{152}\) and ultrasonography are generally employed for initial diagnosis, followed by endoscopic ultrasonography for confirmation. In the majority of cases, PDAC arises in the head region of the pancreas \(^{153}\).

By the time of diagnosis, it is thought that cancer cells may have disseminated to other
sites at the body especially the lungs or liver and thus aggressive systemic chemotherapy is commonly used in order to reduce growth of the primary tumours and invasive, potentially metastatic cells. In cases where no metastasis can be detected (approximately 10%), pancreaticoduodenectomy maybe performed, however, various surgical procedures have been adopted to distinct tumour localisations and stages. Pancreaticoduodenectomy increases the 5-year survival rate to approximately 20%, however long-term survival remains staggeringly low at less than 2%. Moreover, PDAC is strikingly resistant to both chemotherapy and radiotherapy, making complete remission of such cancer extremely rare. Therefore in most cases, patient care focuses primarily on palliation.

There are few established risk factors to develop PDAC, such such smoking, alcohol or coffee consumption, overweight while Vitamin C intake maybe protective. Otherwise there are no diets or supplementations available to protect from PDAC and also life style modification is limited to weight reduction. However apart from these factors, there are a number of other confounding factors which increase the risk for PDAC such as diabetes, and chronic pancreatitis. In particular, certain genetic variants dramatically increase the PDAC risk in hereditary chronic pancreatitis such as mutations in the kationic trypsinogen gene. Underlying genetic alterations have also been identified to give rise to inherited forms of the disease known as Familial pancreatic cancer (FPC). However, such cases are rare (approximately 10%) and of these, only a small fraction have been linked to causal gene mutations or known syndromes. Such families have an estimated 2- to 4-fold increased risk, which dramatically rises to 57-fold when 3 or more relatives are affected. Other genetic conditions in which the incidence of PDAC is increased include mutations in hereditary breast and ovarian cancer associated with BRCA1/BRCA2 (5%), Peutz-Jeghers syndrome (36%) and cystic fibrosis (<5%).
5.6.1. Disease Pathology

There has been much debate as to defining the cell of origin for pancreatic cancer\textsuperscript{170,171}. However, the most widely accepted model of PDAC progression proposes that accumulation of multiple genetic alterations induces changes in acinar cells such that they give rise to well-defined neoplastic duct–like precursor lesions that eventually culminate in a tumour\textsuperscript{172}. The initial transformation event is thought to occur through acinar-to-ductal metaplasia (ADM), a reprogramming event that induces transdifferentiation to a duct-like phenotype\textsuperscript{173}. In the context of additional oncogenic manipulations, ADMs may further differentiate into neoplastic duct-like cells that form lesions and contribute to development of pancreatic cancer. These precursor lesions originate in the epithelium and are termed pancreatic intraepithelial neoplasias (PanINs) and are usually limited to less than 5 mm in size\textsuperscript{172}. PanINs are often located in the pancreatic parenchyma neighbouring infiltrating adenocarcinomas and histologically, show a high degree of morphological variations relative to normal ducts as described in figure 6 and figure 7. Thus they are thought to represent various graded stages of increasingly dysplastic growth. Other premalignant lesions of the pancreas are less well characterised, include intra-pancreatic mucinous neoplasia and mucinous cystic neoplasia\textsuperscript{174,175}. However in the context of the current study, precursor lesions will be discussed in the context of ADMs and PanINs, unless specifically stated otherwise. In recent years, PanINs have undergone stringent classification in mice and men, based on the degree of cytological and architectural atypia and as a result have now been sub-classified into 4 separate lesion types as described in\textsuperscript{172,176,144}.

![Figure 6. Representative images of different grades of PanIN lesions leading to adenocarcinoma. Examples of Normal ducts, PanIN1A/B, PanIN2, PanIN3 and adenocarcinoma are shown from left to right. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Cancer\textsuperscript{144} copyright 2002.](image-url)
After progressing from these precursor lesions, pancreatic cancer is classified by stages, based on resectability in relation to helical CT. The current standard is set according to the American Joint Committee on Cancer tumor–node–metastasis classification standards 2010. This is based on 3 main criteria, designated: Tumour size, Nodal metastasis and distal organ Metastasis (TNM). Once the T, N, and M
groupings have been assessed, this information is combined to assign an overall stage of 0, I, II, III, or IV sometimes combined with a letter in a process known as “stage grouping” as outlined in table 3 below. The stage of cancer is associated with patient survival as demonstrated in a study by Bilimoria et al., which showed that patients with stage 1A cancer have a median survival of approximately 24.1 months, while those with stage IV cancer have a median survival of 4.5 months. Another aspect of tumour classification is based on histological grading. This system classifies cancers on the basis of how abnormal the cells appear from a scale of G1 to G4, where G1 cancers are associated with the most normal looking cell types and the best outlook and is sometimes used in combination with stage grouping.
Table 3. Stage grouping of PDAC according to American Joint Committee on Cancer tumor–node–metastasis classification standards 2010.  

<table>
<thead>
<tr>
<th>Stage Grouping</th>
<th>Tumour Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0</td>
<td>The tumour is confined to the top layers of pancreatic duct cells and has not invaded outside of the pancreas.</td>
</tr>
<tr>
<td></td>
<td>(Tis, N0, M0)</td>
</tr>
<tr>
<td>Stage IA</td>
<td>The tumour is confined to the pancreas and is 2 cm across or smaller (T1).</td>
</tr>
<tr>
<td></td>
<td>(T1, N0, M0)</td>
</tr>
<tr>
<td>Stage IB</td>
<td>The tumour is confined to the pancreas and is larger than 2 cm across (T2).</td>
</tr>
<tr>
<td></td>
<td>(T2, N0, M0)</td>
</tr>
<tr>
<td>Stage IIA</td>
<td>The tumour is growing outside the pancreas but not into major blood vessels or nerves (T3).</td>
</tr>
<tr>
<td></td>
<td>(T3, N0, M0)</td>
</tr>
<tr>
<td>Stage IIB</td>
<td>The tumour is either confined to the pancreas or growing outside the pancreas but not into major blood vessels or nerves (T1-T3).</td>
</tr>
<tr>
<td></td>
<td>(T1-3, N1, M0)</td>
</tr>
<tr>
<td>Stage III</td>
<td>The tumour is growing outside the pancreas into nearby major blood vessels or nerves (T4).</td>
</tr>
<tr>
<td></td>
<td>(T4, Any N, M0)</td>
</tr>
<tr>
<td>Stage IV</td>
<td>The cancer has spread to distant sites (M1).</td>
</tr>
<tr>
<td></td>
<td>Any T, Any N, M1)</td>
</tr>
</tbody>
</table>

In addition to the epithelial components of the main tumour, there is a strong infiltration of stromal cells or fibroblasts around these precursor lesions which continue to expand and are thought to create a microenvironment that supports tumour growth. This is known as the peritumoral desmoplasia and it affects intra-tumoral drug delivery. However two recent studies showed that these large deposits of stroma surrounding some pancreatic cancers may actually have protective properties and that its depletion gives rise to more aggressive tumours. Thus more studies are needed to further understand the contribution of the stroma at different points during tumour formation and
using different models \textsuperscript{185}. Interestingly, a landmark study reported the sub-classification of PDA into three groups: classical, quasi-mesenchymal, and exocrine-like \textsuperscript{186}. Using a combination of transcriptional profiles of primary PDA samples and human and mouse PDA cell lines, they provide evidence for clinical outcome and therapeutic response differences between the different sub-types and have generated defined gene signatures, which may help in stratifying patients for treatment \textsuperscript{186}.

5.6.2. Common mutations associated with PDAC

Along with striking changes in the architecture of the pancreatic epithelium, PanINs and PDAC progression are associated with an anomaly of genetic alterations \textsuperscript{187}. The emergence of genome-wide sequencing studies in recent years has been a major source of new data into the biology occurring at different stages of PDAC and such meta-analysis is helping to build a profile of the genetic landscape in pancreatic cancer. Interestingly, many of the genetic mutations that occur in PDAC have now been identified to be already present at the different PanIN stages, thus establishing the linear progression model of PDAC \textsuperscript{177,188}. Although many of these genetic mutations in chronologic order, there are variances in the exact timing of each mutation between different patients and not all tumours exhibit all genetic abnormalities. This was highlighted in a recent study by Jones et al., which provided comprehensive genetic analysis of 24 individual pancreatic cancers. Here the authors identified an average of 63 genetic abnormalities per tumour, which could be classified into 12 functionally related pathways \textsuperscript{189}. However this study emphasized how complex and heterogenous pancreatic tumours are, as not all tumours had mutations in all pathways and the precise mutations were variable in each pathway. Most recently, genome wide exome sequencing studies further supported PDAC heterogeneity and classified various cases according to frequency and distribution of structural rearrangements. There are also focal amplifications containing potentially druggable oncogenes, inactivation of DNA maintenance genes and DNA damage repair deficiency as well as multiple mutated genes that are prognostic or targetable \textsuperscript{190-193}. Thus, there is a growing list of mutations associated with PDAC onset and progression. However, there are a number of
mutations that are over-represented in PDAC such as Kras, CDKN2A, TP53, BRCA2 and SMAD4/DPC4, some of which will be discussed in detail below.

5.6.2.1. Kirsten rat sarcoma viral oncogene homolog (KRAS).

Kras mutations are thought to be one of the earliest genetic events to occur in pancreatic malignant transition and increase in frequency with disease progression\textsuperscript{194}. Approximately 36%, 44%, and 87% of cancer-associated PanIN-1A, PanIN 1B, and PanIN-2/3 lesions display activating Kras mutations, which increases to between 90-95% of frank PDAC cases\textsuperscript{195}. This knowledge had led to the use of activating Kras mutations in murine models of PDAC, which will be discussed later.

The Kras gene is a guanine nucleotide transferase that belongs to the Ras family of oncogenes, including Neuroblastoma (N) -Ras and Harvey (H)-Ras. The sequences of N-ras and Kras were originally identified in the rat genome in 1981 and subsequently homologues were found in the murine and human genomes\textsuperscript{196}. By 1983, the third member of the mammalian family of ras-related genes, N-ras, had been cloned from neuroblastoma and leukemia cell lines isolated from human tumours\textsuperscript{197}. At the amino acid sequence level, the three different Ras proteins share approximately 85% homology\textsuperscript{196} and it is now apparent that Ras proteins have a normal function as nucleotide-driven switches that communicate extracellular cues to cytoplasmic signalling cascades. However, as these oncogenes normally reside in the genome of normal cells, acquired mutations within particular amino acids via retroviruses causes these proto-oncogenes to become active\textsuperscript{177,194}. The Kras gene is located at position 12p12.1, and directs the production of K-Ras protein. It is a GTPase and thus acts to convert active GTP to the inactive form - GDP. By acting as a regulator of this process, K-Ras integrates signalling from cell surface receptors to a variety of intracellular signalling pathways such as the Mitogen Activated Protein Kinase pathway (MAPK), as phosphoinositide 3-kinase (PI3K) pathway amongst others. Thus, in the normal cellular context, K-Ras plays a role in numerous important aspects of cellular homeostasis, including cell division, differentiation, motility, survival and apoptosis\textsuperscript{198} as outlined in Figure 8.
Figure 8. Ras proteins function as nucleotide-driven switches that relay extracellular cues to cytoplasmic signalling cascades. The binding of GTP to Ras proteins locks them in their active states, which enables high affinity interactions with downstream targets that are called effectors. Subsequently, a slow intrinsic GTPase activity cleaves off the \( \gamma \)-phosphate, leading to Ras functional inactivation and thus the termination of signalling. This on–off cycle is tightly controlled by GTPase-activating proteins (GAPs) and guanine-nucleotide exchange factors (GEFs). GAPs enhance the intrinsic GTPase activity and hence negatively regulate Ras protein function. Conversely, GEFs (also known as GTP-releasing proteins/factors, termed GRPs or GRFs) catalyse nucleotide ejection and therefore facilitate GTP binding and protein activation. Adapted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology \(^{198}\), copyright (2008).

Ras proteins function as nucleotide-driven switches that relay extracellular cues to cytoplasmic signalling cascades \(^{198}\). However, during oncogenic progression, alteration of the normal Kras structure occurs as revealed via sequencing analysis of Kras from different human malignancies. In general, single point mutations resulting from amino acid substitutions were found to affect amino acid residue 12 and with less frequency the position 13 and 61. The incorporation of a different amino acid, usually aspartate at codon 12, induces larger side chains at the GDP/GTP protein binding pocket \(^{199}\) and thus alter the ability of Kras to induce GTP hydrolysis, resulting in an elongated Kras activity as effectively shown by recently by \(^{200}\). Previously it had been postulated that such point mutations give rise to “constitutively active” Kras signalling. However this study provided evidence that such mutations provide an advantage to Kras activation by upstream stimulants and that there is a subsequent delay in the kinetics of reducing Kras activity. Interestingly, a recent study has suggested that both wildtype H-ras and N-
ras may contribute to mutant Kras-driven tumorigenesis \(^{201}\). Although activation of Kras\(^{G12D}\) or Kras\(^{G12V}\) in the murine pancreas is sufficient to initiate development of ADM, and PanIN lesions, there is a very low frequency and long latency to invasive metastatic PDAC required, thus suggesting that additional genetic aberration are required for disease progression \(^{177}\).

5.6.2.2. **The TP53 tumor suppressor**

The TP53 tumour-suppressor gene is inactivated in approximately 50 – 75% of pancreatic cancers \(^{202,203}\). Generally such mutations arise in advanced lesions (PanIN-3 or later) and are primarily due to intragenic missense mutation of the DNA binding domain coupled with loss of the second allele \(^{204,205}\). TP53 is a master regulator of the DNA damage and cellular stress response, and controls a large number of proteins involved in cell cycle arrest, senescence and apoptosis \(^{206}\). Upon loss of tp53, cells lose their ability to control such mechanisms effectively, resulting in accumulation of genetic abnormalities in cells that continue to survive and divide \(^{207}\).

5.6.2.3. **SMAD4**

Another recurrent alteration in PDAC is associated with loss of the SMAD4/DPC4 (Mothers against decapentaplegic homolog 4 / deleted in pancreatic carcinoma 4) gene. Located on chromosome 18q21, this gene is inactivated in approximately 55% of PDAC cases, with the first mutations occurring around the PanIN 3 stage. Such mutations generally occur through homozygous deletion (30%) or via intragenic mutations with loss of the second allele (25%) \(^{208,209}\). SMAD4 is a key regulator of the TGF-\(\beta\) signalling cascade, which is a potent inhibitor of epithelial cell survival and growth. TGF-\(\beta\) is a cytokine which signals through serine/threonine kinase receptors and in normal, healthy epithelial cells, acts as an inhibitor of epithelial cell growth and survival \(^{208}\). In cancer however, TGF-\(\beta\) can promote the growth of malignant epithelial tumours via enhanced cell proliferation, motility and EMT \(^{210}\).
5.6.3. Genetically engineered mouse models (GEMMs)

In recent years, GEMMs have been used in order to mimic common genetic mutations that occur in human malignancies. The advantage of such models is that the mutations may be introduced either constitutively or conditionally in a tissue specific context. Moreover, this technology allows for manipulation of mutant genes that may represent gain-of-function oncogenes or loss-of-function tumour suppressor molecules and thus, the precise contribution of a variety of related or diverse pathways can be interrogated in disease initiation and/or progression. The efficacy and advantages of such models is highlighted by the rapid development of such models for a variety of common malignancies including breast, lung, colon, prostate and pancreatic cancers. 

The development of the first GEMM of PDAC utilized a mice strain with a knock-in of the endogenous Kras allele, whereby the amino acid glycine (G) at position 12 is substituted with an aspartic acid (D) and this is preceded with a floxed STOP transcriptional cassette (Lox-Stop-Lox). These mice were then crossed with transgenic mice harbouring the bacterial Cre recombinase under the control of the pancreas-specific promoters Pdx1 or Ptf1a/P48. The Cre-lox system is a genetic tool that allows for site-specific recombination events to occur in genomic DNA. In order to integrate the Cre-Lox P technology into mice, generally, Cre and Lox P mouse strains are developed separately and finally crossed in order to produce the final target strain containing both Cre and Lox P. The offspring were designated as either Pdx1 Cre; LSL-KRasG12D or Ptf1a Cre; LSL-KRasG12D. Upon successful recombination whereby the STOP transcriptional cassette is removed, one allele of the new strain will express the mutated KrasG12D product, which results in constitutively active Kras signalling. Of note, both LSL-KRasG12D and Pdx1-/- Ptf1a- Cre may only be bred as heterozygotes as homozygous mice are embryonic lethal. In addition, these mice are commonly termed as “KC mice”, according to the current nomenclature. Both of these models recapitulate the full spectrum of changes observed in human patients with PDAC, with complete penetrance and will be described in detail below. Notably such an approach has now been adapted to allow for a timed controlled onset of recombination by utilizing a tamoxifen-inducible Cre system. Moreover, a new GEMM of PDAC has been reported recently and has some benefits over conventional Cre-loxP-based models. By
combining flippase-FRT and Cre-loxP recombination strategies, this study developed an inducible dual-recombinase system, which allows for the controlled investigation of multistep carcinogenesis. Importantly, this technology allows one to genetically manipulate different subpopulations and for selective targeting of the tumour microenvironment. Interestingly, a study has shown that the acinar differentiation factor PTF1A inhibits initiation of PDAC and thus, the implications having a heterozygous mutation as used in the p48Cre GEMM of PDAC remain to be clarified.

5.6.1. Signalling pathways associated with PDAC

Interestingly, many of the signalling pathways that contribute to early embryonic development and organogenesis become re-activated during carcinogenesis including the Wnt/β-catenin, EGF, TGF and FGF signalling pathways. In particular, GEMMs have shown that mis-expression of members of these signalling families contribute to disease initiation and/or progression.

TGF beta signalling comprises a diverse group of proteins and signalling molecules that play key roles in pancreatic development such as Activin A and Nodal. Moreover, BMPs are prominently expressed in the developing pancreas and regulate cellular differentiation and cell growth. Notably, alterations within the TGF signalling family are frequently associated with PDAC and the loss of TGF signalling members suggests that it may act as a tumour suppressor as highlighted by the loss of SMAD4 in a GEMM of PDAC. However increased expression of other family members such as BMPs has been reported, indicating that a complex contribution of several TGF beta family members to disease progression may exist.

The FGF family represents a group of 14 heparin-binding proteins that have numerous roles in development including endoderm specification. This family of proteins act through one of four FGF receptors (FGFRs) and activate a variety of downstream signalling cascades including the phosphatidylinositol-3 kinase (PI3K) and the Ras-Raf1-MEK-MAPK pathways. Perturbed FGF signalling is associated with PDAC and interestingly there is a correlation with FGF expression levels, tumour stage and patient survival.
The EGF family includes more than 30 different factors that serve a variety of important cellular processes including cell proliferation and differentiation, as well as EMT. They exert their actions through a number of ErbB tyrosine kinase receptors and remarkably, many different EGF ligands and receptors are overexpressed in PDAC. TGF-α is one such ligand and has been effectively shown to induce acinar-to-ductal metaplasia, leading to early PanIN lesion progression.

Wnt signalling is required for endodermal patterning and pancreatic organogenesis. Soluble Wnt ligands bind and activate the Frizzled family of receptors causing translocation of β-catenin to the nucleus and subsequent activation of downstream transcription factors. Abnormal Wnt signalling has been identified in a large number of PDACs and has been shown to be detectable already in PanIN stage lesions and to promote tumour progression.

Taken together, this non-exhaustive list demonstrates that many signalling pathways important for early embryonic development and specification of the endodermal and pancreatic lineages, become re-activated and during pancreatic cancer progression. Genetic aberrations within these molecular pathways are implicated with disease burden and thus more research is needed to understand the contribution of such pathways at different points of precursor lesion formation, tumour growth and metastasis.

**5.6.2. Pancreatic Cancer Stem Cells.**

It is well established that tumours comprise heterogeneous entities that contain a variety of different cell populations. In recent years, there has been a significant effort to identify the “cell of origin” in various solid tumours. This has led to the identification that only a minor subpopulation of cells within a tumour are responsible for initiation of tumour growth and subsequent expansion. Such cells have been termed “cancer stem cells” (CSCs) or tumour initiating cells (TICs) due to their self-renewal potential and multipotent capability. The cancer stem cell model is supported by the reports whereby such cells have been shown to give rise to a progeny of various differentiation states that can colonize the tumour and recapitulate the heterogeneity of the primary tumour. The basis for such a model comes from studies of leukemia stem cells. Here the authors could demonstrate that primitive normal stem cells were able to differentiate in...
vivo into leukemic blasts, indicating that the disease-causing cell originates through differentiation from a hierarchy. Since then, CSCs have been identified in numerous tissues and isolated from a variety of solid tumours including brain, colon and breast, lung and prostate tumours. 

In the pancreas, CSCs were originally reported in a study in 2007. Using cell surface markers, this group identified a subpopulation of cells with the marker combination of CD24, CD44 and ESA. These cells were extremely rare (0.2 – 0.8 %) in relation to the total pancreatic tumour cells isolated and had the ability to self-renew and further differentiate into an array of different cell types, recapitulating the heterogeneity of the primary tumour. Subsequently, another study showed that CD133 expression could be used to mark cells with enhanced proliferative capacity and tumour initiating potential, which were highly resistant to chemotherapy. Moreover, at the invasive front of pancreatic tumours, some CD133 positive cells also expressed CXCR4 and represent cells with metastatic potential.

An interesting point between these studies and the marker expression profiles used to identify such cells is that only approximately 14% overlap exists between CD44+CD24+ESA+ and CD133+ cells. Thus more work is needed to better specify which marker combination may be useful in order to target the cancer stem cell population and in order to generate a precise molecular profile of these cells. The origin of CSCs in the respective tissues remains to be determined as it is unclear if they represent a population of cells that reside in quiescent state in the normal adult tissue or if a de-differentiation occurs within particular cell types.

5.6.3. Novel mutations to define PDAC subgroups

Mutations in genes including KRAS, P16/CDKN2A, TP53 and SMAD4/DPC4 genes are commonly denoted as “driver mutations” in PDAC, as they have been shown to enhance neoplastic transformation and tumour formation, as underlined by the use of mouse models. However, in recent years, numerous studies have identified new mutations that are implicated with PDAC. Thus it is important to understand the contribution of different genetic abnormalities to disease risk and tumour progression, and in effect, distinguish between “driver” mutations and “passenger” mutations which
maybe relevant to disease prevalence or disease phenotype, or which maybe neutral and not significantly contribute to the disease. In particular, with the advent of next generation sequencing studies, a range of previously un-appreciated mutations implicated with PDAC have been reported, thus providing a detailed molecular characterization of PDAC and associated precursor lesions.

A study by Carter et al. used the analytical technique of Cancer-specific High-throughput Annotation of Somatic Mutations (CHASM) and studied 963 somatic missense mutations discovered in 24 PDACs. They identified new driver mutations, of which 4 occurred in kinases including PIK3CG, DGKA, STK33 and PRKCG, which have the potential to integrate with the Kras signalling pathway and likely affect tumour cell properties including cell growth, survival, migration, invasion and metastatic potential. Notably, genome-wide exome sequencing studies of PDACs have garnered a lot of attention in recent years and have revealed numerous variations in chromosomal structures and DNA maintenance genes. Many of the mutated genes have either prognostic significance or are targetable. The correlation of mRNA expression, DNA copy number variations and miRNA levels in a survival-based molecular array analysis of early stage PDAC has revealed molecular alterations associated with an unfavourable prognosis. In particular, mutations have been highlighted in genes such as BRCA1, ATM, EPC1, ARID2, KDM6A and PREX2 in subsets of PDAC, as outlined in figure 9. The contribution of such mutations to disease risk and pathophysiology is unclear, although notably, some of these genes have been implicated in other cancers. Interestingly, it has been reported that up to 10% of PDACs maybe due to genetic susceptibility, whereby families in which two or more first-degree relatives are affected is referred to FPC. Such genetic factors are associated with increased risk of cancer and varying penetrance and degree for each mutation. Therefore, as more sequencing data is produced and more genes are implicated in the disease, it is important to verify the contribution of such genes and genetic modifications to disease penetrance and disease pathophysiology. Thus, a thorough characterization of novel mutations should provide a system to distinguish “driver” mutations from “passenger” mutations that ultimately characterise each individual cancer and may allow for stratification of individuals with increased risk of
PDAC, as well as provide opportunities for early detection and specific treatment strategies.

**Figure 9.** Common and rare mutations in PDAC with therapeutic and or prognostic relevance, based on data from 190-191.

<table>
<thead>
<tr>
<th>GENE CLUSTER</th>
<th>Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreatic cancer genes</td>
<td>KRAS, TP53, CDKN2A, SMAD4, MAP2K4, TGFBR2, ACVR1B, RNF43, BRAF, SF3B1, PREX2</td>
</tr>
<tr>
<td>SWI/SNF mediated chromatin remodelling</td>
<td>KDM6A, ARID1A, ARID1B, PBRM1, SMARCA2, SMARCA4, MLL2</td>
</tr>
<tr>
<td>DNA damage repair</td>
<td>BRCA2, BRCA1, PALB2, RPA1, ATM, STK11, MLH1, MSH2, MSH6, POLE, EXO1, PMS2, BCLAF1, FANCA, FANCF, FANCM</td>
</tr>
<tr>
<td>Axon guidance</td>
<td>ROBO1, ROBO2, SLIT2</td>
</tr>
<tr>
<td>Other known oncogenes</td>
<td>MYC, GATA6, CDK6, NOV, MET, SOX9, ERBB2, PIK3CA, PIK3R3, FGFR2, APC, AXIN1, GNAS, CDK4, TGFBI</td>
</tr>
</tbody>
</table>

**Pathways:**
- RB pathway
- TGFβ pathway
- ERBB2/HER2 pathway
- Wnt signaling pathway
- FGF signaling pathway
### 5.7. Ataxia-telangiectasia mutated (ATM)

Recent studies have identified that ATM is among the 16 most commonly mutated genes in human PDAC \cite{189,193,248} and notably, mutations can be found in precursor lesions. Interestingly, ATM was mutated in the germline of sub-sets of PDAC patients \cite{250,255}. It was previously described that ATM heterozygosity co-operates with loss of Brca1 in anaplastic breast cancer \cite{256} and one sequencing analysis of human PDAC cases linked ATM and BRCA mutations in genetically unstable PDAC subtypes \cite{190}. Moreover, one recent study correlated ATM protein kinase levels with survival analysis and showed that low protein levels of ATM stratify patients into a group associated with poor survival \cite{257}.

ATM is a member of the phosphatidylinositol 3 kinase-like kinase (PIKK) family of Serine/Threonine-protein kinases. Its name derives from the product of the gene that is mutated in the human genetic disorder ataxia telangiectasia (AT). AT is a rare autosomal recessive disease occurring in early in childhood. The incidence varies from 1 out of 40,000 to 1 out of 100,000 new births and the carrier frequency is approximately 1\% of the population \cite{258}. The disease is caused by a deficiency of the ATM protein kinase, primarily due to nonsense mutations (85\%) that lead to unstable ATM protein variants, while less than 15\% arise due to missense mutations. Several hundred different ATM mutations have been identified in AT patients, most of which occur as heterozygous. AT is characterized by severe immunodeficiency, progressive neurodegeneration (cerebellar degeneration and progressive ataxia) as well as increased incidence of cancer, particularly lymphoid tumourgenesis. Moreover, AT patient’s exhibit hypersensitivity to ionizing radiation, have higher indices of metabolic disease including increased susceptibility to diabetes and impaired glucose metabolism as well as infertility \cite{259}. The associated phenotype of AT is thought to be due to defects in the DNA damage and repair pathways that give rise to genomic instability as well as escape from normal cell cycle control due to defective ATM signalling. ATM knockout mice recapitulate much of the phenotype of AT patients, except for the absence of progressive neurodegeneration. In contrast, mice with a kinase-dead ATM (catalytically inactive) \cite{260-262} are embryonic lethal, indicative of a much more severe phenotype \cite{263,264}. The disparity for this phenomenon is unclear, however it has been shown that the
kinase-dead mutant was recruited to DSB sites and thus its presence is thought to disrupt the rest of the DNA damage response network.

5.7.1. ATM protein kinase structure:
ATM is a large protein kinase of 370 kDa located on human chromosome 11q22-23. It is evolutionary conserved as homologs have been found in all eukaryotes and mouse ATM, which is encoded on chromosome 9 shares 84% sequence homology to the human protein. As a member of the PIKK family, ATM contains the common structural motifs including a conserved kinase domain (KD), a FRAP-ATM-TRRAP (FAT) domain, and a FAT-C terminal (FATC) domain near the C-terminus as outlined in figure 10. In addition the protein also contains N-terminal HEAT repeats (named after the proteins huntingtin, elongation factor 3, the A subunit of PP2A and TOR1) and a substrate-binding domain. Notably, much of the domain structure has been determined through comparison of related phosphatidylinositol-like superfamily proteins, in vitro studies and also due to extensive mapping of patient mutations, which have highlighted numerous post-translational modification sites. Due to the large complex nature of the ATM protein, a detailed x-ray structure is still elusive. However it is known that in its inactive form, ATM exists non-covalently linked as a dimer or in a multimeric conformation. Although ATM is traditionally considered as a tumour suppressor gene, more recent studies have identified a diverse array of signalling pathways that integrate with ATM and suggest a broader role for ATM in intracellular signalling mechanisms. ATM activation can essentially be divided into two classifications: the canonical pathway in which the DNA damage checkpoint becomes activated and several non-canonical modules which are stimulated by various forms of cellular stress.
5.7.2. Canonical ATM signalling

In response to DNA double strand breaks (DSB), the ATM dimer dissociates and the active monomers are recruited to the specific DNA DSB sites \(^{266}\). Activation of ATM due to DBSs is an intricate process that involves the Mre11-Rad50-NBS1 (MRN) complex. It is known that autophosphorylation of ATM occurs in at least four specific residues (S367, S1893, S1981, and S2996) through interactions with several opposing phosphatases such as protein phosphatase2A (PP2A), protein phosphatase5 (PP5) and wild type p53-induced phosphatase1 (WIP1). Notably, in site-directed mutagenesis studies, inhibition of these residues in mouse or human ATM-deficient cells, human cells remained defective in their ATM-dependent DNA damage response, while in murine cells abrogation of these autophosphorylation sites did not affect the ATM DDR functions \(^{269,270}\). Thus the precise role of autophosphorylation in activation of ATM will require systematic and detailed analysis in a species-specific context. Although the exact role of the MRN complex remains unclear in the monomerization of ATM dimers, it appears that the MRN complex certainly plays a role in recruitment of ATM to the site of DSBs, in particular the ubiquitination of the NBS1 subunit, which interacts directly with ATM at the C-terminus. Various studies have highlighted that ATM is recruited to DNA damage at diverse genomic locations, in particular in heterochromatin \(^{271}\). The retention
of ATM at DSBs has an absolute requirement for its kinase activity as a kinase-dead mutant of ATM was shown to be ineffective in forming damage-induced foci \(^{272,273}\).

### 5.7.3. Non-canonical ATM signalling

Accumulating evidence suggest that ATM is a complex and versatile protein kinase that integrates with a large number of signalling pathways in order to maintain cellular homeostasis \(^{267}\). The high number of predicted downstream targets places ATM as a signalling core and the precise actions of ATM is subject to the complex actions of cofactors, substrates and feedback regulators, in a cell specific manner \(^{266}\). These include activation in response to oxidative stress, as a redox sensor and in response to insulin signalling \(^{274}\) as outlined in **Figure 11**.

**Figure 11. ATM interaction with metabolic signalling pathways.** ATM is activated in response to Ionizing Radiation (IR), insulin signalling and Reactive Oxygen Species (ROS). Signalling pathways initiated by IR, insulin, and ROS are indicated by black, blue, and red arrows, respectively. Reprinted from Trends Biochem Sci. 37(1): 15–22, Ditch S. and Paull TT. The ATM protein kinase and cellular redox signaling: beyond the DNA damage response, Copyright (2012) with permission form Elsevier \(^{266}\).
5.7.4. Germline mutations and somatic mutations in ATM:
The first proposal that relatives of patients with A-T might be at increased risk of cancer was recorded more than 20 years ago \(^{275}\). This study suggested the relative risk of cancer was 2.3 for men and 3.1 for women in 110 A-T families, where breast cancer was the most strongly associated. While there has been a long established link between A-T patients and cancer predisposition, different mutations within the gene can interrupt a variety of pathways giving rise to a range of diverse phenotypes. According to the National Institutes of Health America, the risk of an A-T patient developing any cancer is 37-fold higher than individuals in the general population, while the risk of developing lymphoid tumours is 100-fold higher than in the general population. More than 300 ATM mutations have been identified to date in A-T patients \(^{51,276,277}\). Despite this extensive variability, the majority of such mutations lead to truncation of the ATM protein often with large deletions. Interestingly, the truncated ATM derivatives are very unstable and become degraded, leaving no protein product detectable. Less deleterious variants of ATM mutations exist and usually involve small in-frame deletions or splicing mutations, which over time lead to considerably reduced, but detectable ATM proteins levels \(^{277}\). Therefore, such mutation profiles suggest that in A-T patients, the phenotype is defined by null ATM-alleles. Remarkably, recent studies have suggested that between 0.5 % and 1 % of the general population may be heterozygous for ATM mutations \(^{258,275,278}\) and such mutations are also associated with an increased risk of cancer. In particular, breast cancer has been reported to be fivefold more likely to occur in females with heterozygous ATM mutations when under the age of 50 years old \(^{256,278-280}\).

Unlike germline mutations, somatic mutations refer to change in the genetic structure that are neither inherited nor transferred to offspring. Such alterations may give rise to cancer or other diseases or may not have any detrimental affects at all \(^{281}\). In relation to cancer, somatic ATM mutations such as deletions or inactivating mutations have also been identified in numerous cancer types including chronic lymphocytic leukaemia patients, (~5%) \(^{282}\), lung cancers (~8%) \(^{283}\), breast cancer (~55%) \(^{284}\). In many cases, such mutations often occur in a mutual exclusive manner with other mutations in tumour suppressor genes such as TP53 \(^{282}\).
6. Hypothesis and Aims

6.1. The role of Tbx3 at the transition from pluripotency to lineage commitment

Pluripotent ESCs are characterized by their capacity for unlimited self-renewal and the potential to differentiate into every cell / tissue of an organism without giving rise to a complete organism on their own. It is well established that there is a key network of transcription factors and signalling molecules required for the maintenance of the self-renewal phenotype including OCT4, Sox2 and LIF. However in order to escape from the pluripotent state, the levels of such factors must fluctuate in precise manners in order to allow for subsequent differentiation and lineage commitment. Tbx3 has been shown to interact with the core pluripotency gene network and it has been reported that Tbx3 expression was necessary to maintain cells in a pluripotent state.

Therefore the primary goals of this project were:
1. To characterize the expression of Tbx3 in mESCs at the exit from pluripotency and the onset of differentiation in vivo and in vitro.
2. To provide a mechanistic insight into how Tbx3 may direct cell lineage commitment.

6.2. A protocol to generate Definitive Endoderm from human iPSCs

Mammalian development is a rapid and dynamic process, which sees the embryo undergo a tightly regulated series of morphological and physiological changes to allow for a fully formed organism. iPSCs derived from healthy donors allows for unparalleled potential to accurately model normal development. In addition, iPSCs derived from diseased patients provide an effective system to study a specific disease and to investigate the underlying phenotype. However, a pre-requisite for performing such studies are standardized and reproducible differentiation protocols to specific cell / tissue types such as endoderm, which in turn gives rise to internal organs such as the digestive tract, liver, gallbladder, and pancreas.
In this study, we aimed to:

1. Generate a robust and convenient protocol to differentiate iPSCs to DE.
2. Characterize the expression of a specific set of ion channels during the different stages of differentiation.

6.3. The role of ATM in PDAC using a GEMM

Pancreatic cancer is a complex disease in which multiple signalling pathways are disturbed. Recent data from human PDAC sequencing studies suggest that the ATM kinase is implicated in PDAC progression. Thereby, the generation of a conditional loss of function model for ATM in a PDAC GEMM setting should provide a valuable tool to address the mechanisms by which ATM plays a role in PDAC initiation and/or progression.

The specific objectives of the study were:

1. To establish a mouse model for loss of ATM activity in pancreatic cancer and characterise the phenotype of the mice.
2. To analyze the mechanisms by which loss of ATM may influence initiation and progression of PDAC pathophysiology.
7. Bibliography


Han, J. & Sidhu, K. S. in Methodological Advances in the Culture, Manipulation and Utilization of Embryonic Stem Cells for Basic and Practical Applications 15. Reprogramming Somatic Cells by Fusion with Embryonic Stem Cells: Present Status and Prospects in Regenerative Medicine, Ch. Reprogramming Somatic Cells by Fusion with Embryonic Stem Cells: Present Status and Prospects in Regenerative Medicine, (InTech, 2011).


stem cells.


Kartikasari, A. E. et al. The histone demethylase Jmd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation. EMBO J 32, 1393-1408, doi:10.1038/emboj.2013.78 (2013).


Spring, K. et al. Mice heterozygous for mutation in Atm, the gene involved in ataxia-telangiectasia, have heightened susceptibility to cancer. Nat Genet 32, 185-190, doi:10.1038/ng958 (2002).


8. Appendix

8.1. Tbx3 directs cell-fate decisions toward mesendoderm.

8.2. Definitive endoderm formation form plucked human hair-derived induced pluripotent stem cells and SK channel regulation

8.3. Loss of Ataxia telangiectasia-mutated accelerates pancreatic cancer formation and epithelial-mesenchymal transition.
TBX3 Directs Cell-Fate Decision toward Mesendoderm

Clair E. Weidgang,¹,2,3 Ronan Russell,¹,3,13 Purushothama R. Tata,²,4,13,15 Susanne J. Kühl,⁴ Anett Illing,¹ Martin Müller,¹ Qiong Lin,⁷ Cornelia Brunner,⁵ Tobias M. Boeckers,⁶ Kerstin Bauer,⁸ Apriliana E.R. Kartikasari,⁹ Yanchun Guo,²,4 Melanie Radenz,²,⁴ Christof Bernemann,¹⁰ Matthias Weiß,¹¹ Thomas Seufferlein,¹ Martin Zenke,⁷ Michelina Iacovino,¹²,16 Michael Kyba,¹² Hans R. Schöler,¹⁰ Michael Kühl,⁴ Stefan Liebau,⁶,¹⁴,* and Alexander Kleger¹,14,*

¹Department of Internal Medicine I
²International Graduate School in Molecular Medicine
³Graduate College 1041
⁴Institute for Biochemistry and Molecular Biology
⁵Institute for Physiological Chemistry
⁶Institute for Anatomy & Cell Biology
⁷Ulm University, 89081 Ulm, Germany
⁸Department of Cell Biology, Institute for Biomedical Engineering, Medical Faculty, RWTH Aachen University, 52074 Aachen, Germany
⁹Institute of Molecular Medicine and Max-Planck-Research Group on Stem Cell Aging, 89081 Ulm, Germany
¹⁰Department of Medicine, University of California, Los Angeles, CA 90095, USA
¹¹Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
¹²University Medical Centre, Renal Department, Centre for Clinical Research, 79106 Freiburg, Germany
¹³Department of Pediatrics, Lillehei Endowed Scholar, Lillehei Heart Institute, University of Minnesota, Minneapolis, MN 55455, USA
¹⁴These authors contributed equally to this work
¹⁵These authors contributed equally to this work
¹⁶Present address: Center for Regenerative Medicine, Harvard Stem Cell Institute, Massachusetts General Hospital, Boston, MA 02114, USA
*Correspondence: stefan.liebau@uni-ulm.de, alexander.kleger@uni-ulm.de (A.K.)
http://dx.doi.org/10.1016/j.stemcr.2013.08.002
This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

SUMMARY

Cell-fate decisions and pluripotency are dependent on networks of key transcriptional regulators. Recent reports demonstrated additional functions of pluripotency-associated factors during early lineage commitment. The T-box transcription factor TBX3 has been implicated in regulating embryonic stem cell self-renewal and cardiogenesis. Here, we show that TBX3 is dynamically expressed during specification of the mesendoderm lineages in differentiating embryonic stem cells (ESCs) in vitro and in developing mouse and Xenopus embryos in vivo. Forced TBX3 expression in ESCs promotes mesendoderm specification by directly activating key lineage specification factors and indirectly by enhancing paracrine Nodal/Smad2 signaling. TBX3 loss-of-function analyses in the Xenopus embryo underline its requirement for mesendoderm lineage commitment. Moreover, we uncovered a functional redundancy between TBX3 and Tbx2 during Xenopus gastrulation. Taken together, we define further facets of TBX3 actions and map TBX3 as an upstream regulator of the mesendoderm transcriptional program during gastrulation.

INTRODUCTION

Pluripotent embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs) are characterized by continuous self-renewal while maintaining the potential to differentiate into all three germ layers. The regulatory networks of maintaining ESC pluripotency have been described in great detail (Ying et al., 2008), and, similarly, there is a vast wealth of knowledge on key players that regulate differentiation of pluripotent stem cells. Multiple pathways including WNTs (wingless-related MMTV integration site), transforming growth factor (TGF)-β, BMP (bone morphogenetic protein), and fibroblast growth factor (FGF) signaling act in concert with a combination of key transcription factors to coordinate lineage commitment (Blair et al., 2011).

Recent reports suggest that pluripotency-associated transcription factors engage in additional functions during the early phases of germ layer specification and commitment. OCT4 and NANOG promote mesodermal as well as endodermal fate and limit neuroectoderm differentiation. In contrast, SOX2 enhances neuroectoderm while restricting mesoderm and endoderm development. Accordingly, groups of pluripotency factors were defined as mesendoderm class (e.g., OCT4, Nanog, KLF5) or neuroectoderm class (e.g., SOX2, RBPJ) ESC genes (Thomson et al., 2011).

Preceding mouse gastrulation, the rather symmetrical embryo is prepatterned by regional differences in gene expression and fluctuating levels of signaling pathways along the embryonic axes. Signaling gradients including NODAL and canonical WNT at the posterior pole of the embryo promote the formation of the primitive streak, accompanied by the expression of early differentiation marker genes (Arnold and Robertson, 2009).
As one of the early events during gastrulation, definitive endoderm (DE) and anterior mesoderm derivatives, including cardiovascular and head mesenchyme progenitors, are generated from a transient precursor cell population located in the region of the anterior primitive streak. This cell population is commonly referred to as mesendoderm and is marked by the expression of genes including Chordin (Chrd), Eomesoderm (Eomes), Foxa2, Goosecoid (Gsc), and LIM-homeobox1 (Lhx1) (Tada et al., 2005).

An evolutionarily conserved regulatory network, including NODAL and BMP signaling, controls mesendoderm induction as well as successive specification of DE and mesoderm lineages (Zorn and Wells, 2009). Members of the T-box transcription factor family exert highly specific and crucial functions throughout development (Arnold and Robertson, 2009). During gastrulation EOMES regulates the specification of DE and cardiovascular mesoderm from the anterior primitive streak region, Brachury (T) is required for expansion of the posterior mesoderm and TBX6 specifies intermediate primitive streak derivatives (Pereira et al., 2011). In midgestation murine embryos, TBX3 plays important roles in cardiac chamber specification, hepatic bud expansion, and limb patterning (Chapman et al.,1996; Hoogaars et al., 2007). Heterozygous mutations in the human Tbx3 gene were identified in the Ulnar-mammary syndrome, characterized by limb, genital, apocrine, and cardiac abnormalities (Bamshad et al., 1997). In pluripotent stem cells, TBX3 helps to maintain pluripotency by mediating LIF/STAT signaling (Niwa et al., 2009) and facilitates reprogramming by direct binding and activation of the OCT4 promoter (Han et al., 2010). TBX3 modulates the formation of extraembryonic, visceral endoderm (VE) by directly activating GATA6 expression (Lu et al., 2011) and acts as a downstream activator of WNT signaling (Price et al., 2012). In contrast, TBX3 is one of the transcriptional regulators that is highly enriched in DE progenitor cells (Cheng et al., 2012). Further, TBX3 in concert with the histone demethylase JMJD3 and EOMES are involved in endoderm formation (Kartikasari et al., 2013).

Here, we identified a function for TBX3 in early lineage commitment toward DE and anterior mesoderm derivatives. In murine and Xenopus embryos, TBX3 expression coincides with gastrulation onset, and TBX3 loss of function in Xenopus affects mesendoderm marker gene expression and impairs gastrulation. Our data suggest that TBX3 promotes lineage commitment toward DE and anterior mesoderm derivatives in a dual fashion: first, TBX3 directly regulates key lineage determining transcription factors (cell autonomous). Second, we demonstrate a central role for TBX3 in Nodal-mediated paracrine signaling (non-cell autonomous).

**RESULTS**

**TBX3 Expression Is Dynamically Regulated in Early Embryonic Development**

Published transcriptome data of preimplantation stage mouse embryos (Guo et al., 2010) indicated moderate Tbx3 expression beginning at the 4-cell stage, increasing toward the blastocyst stage (Figure 1A). To determine early embryonic expression of TBX3, we analyzed messenger RNA (mRNA) and protein distribution of TBX3 in pre- and postimplantation stage embryos. Immunofluorescence (IF) analysis demonstrated Tbx3 colocalizing with OCT4 in the ICM of E3.5 blastocysts, whereas protein was absent at the 2-cell stage (Figure 1B; data not shown). mRNA in situ hybridization (ISH) on early pregastrulation stage embryos (E6.25) showed Tbx3 expression in the proximal posterior pole of the epiblast in addition to the reported expression in the extraembryonic structures (Figure 1C, left panel). At midgastrulation stage (E7.5) Tbx3 expression was predominantly found in the extraembryonic VE and in a ring of mesoderm close to the embryonic-extraembryonic intersection (Figure 1C, right panel). At late gastrulation stages (E7.75, E8.25), Tbx3 RNA was observed in the cardiac crescent and the tail region of the embryo (Figure 1D). We found TBX3-reporter expression in E6.5 embryos ( GFP expression driven by a 160 kbp bacterial artificial chromosome with TBX3 and flanking sequences [Horsthuis et al., 2009]) in extraembryonic structures as well as in the proximal posterior epiblast (Figure 1E). We then wondered whether TBX3 expression is conserved in different species and performed semi quantitative RT-PCR and ISH for t bx3 mRNA on different stages of Xenopus embryos. We detected a weak maternal t bx3 expression with a marked increase at the onset of gastrulation (stage 10, Figure 1F). At later gastrulation stages, t bx3 mRNA was readily detectable in early mesodermal cells and in the developing heart region (Figure 1G).

In summary, our results demonstrate that TBX3 is dynamically regulated during early development and expressed at specific sites of gastrulation onset in both mouse and Xenopus embryos.

**TBX3 Promotes ESC Differentiation toward Mesoderm and Endoderm**

During embryoid body (EB) differentiation, we observed moderate Tbx3 mRNA expression in undifferentiated ESCs (Figure 2A) followed by a downregulation during early differentiation (days 1 and 2). Subsequently, expression rapidly increased with a peak on day 3, falling again until day 6. Comparable results were achieved by western blot (WB) and IF (Figures 2B and 2C). Published transcriptome data of several differentiating ESC lines
Figure 1. TBX3 Expression in Developing Mouse and Xenopus Embryos
(A) Tbx3 (blue) in preimplantation embryos plotted with other mesendoderm class ESC genes (Oct4, red; Klf5, brown; Klf9, pink; T, violet). Data were obtained upon reanalysis of published data sets.
(legend continued on next page)
confirm these findings (http://www.fungenes.org). Interestingly, elevated TBX3 levels coincided with the expression of critical lineage determining factors such as EOMES, T, SOX17, and MESP1 in qPCR and IF (Figures 2D–2G). TBX3 expression was found to colocalize with the early differentiation markers EOMES (67% ± 7%), T (46% ± 2%), and SOX17 (51% ± 8%) at day 4 of EB differentiation (Figure 2H). To test whether TBX3 is specifically enhanced in subsets of differentiating ESCs, we used reporter cell lines to trace transcription of T and Sox17, markers of mesoderm and endoderm. After differentiation for 4 days, T- or Sox17- positive cells were fluorescence-activated cell sorting (FACS)-purified, and remaining pluripotent cells were excluded from the analyses using the pluripotency reporter DPPA4-RFP or SSEA1 staining (Figures 2I and 2K; data not shown). qPCR shows enriched Tbx3 expression in cells committed toward the mesoderm or endoderm cell fate (Figures 2J and 2L) and IF confirmed that Sox17 or T positive-sorted cells coexpress Tbx3 (Figures 2M and 2N). These results indicate that endogenous TBX3 expression in differentiating ESCs is correlated with mesendoderm cell-fate specification. A study reported that TBX3 improved iPSC reprogramming and used chromatin immunoprecipitation (ChIP) sequencing to identify genes directly regulated by TBX3 during pluripotent conditions (Han et al., 2010). Reanalyzing these data, we found that TBX3 binding was enriched on the promoters of promoting pluripotency factors including Klf5, Klf4, and Oct4 and promoters of key lineage determining factors such as Sox9, Nkx6.1, and Eomes (Figure S1 available online). This suggests that TBX3 may play an important role during the switch from pluripotency to differentiation toward the mesendodermal lineage.

In order to test this hypothesis, we generated a targeted, inducible Tbx3 expression allele in ESCs to allow for the temporally regulated and dose-dependent expression of Tbx3 (iTbx3 ESCs; Figures S2A–S2D). We used a recombination system in the HPRT locus, where a doxycycline (Dox)-inducible promoter regulates expression of the Tbx3 mRNA (Iacovino et al., 2011). This Tbx3 allele was homogenously induced in all cells upon Dox-exposure. Using EB differentiation, we assessed the effects of TBX3 overexpression on marker genes of all three germ layers (Figure 3A). In Tbx3-overexpressing cells, enhanced differentiation toward mesoderm and endoderm progenitors was confirmed by IF for EOMES, T, and SOX17 (IF and WB) in day 4 EBs (Figure 3B; Figures S2E–S2I) and early marker transcripts of mesendoderm (Chd1, Gsc, Eomes, Lhx1 [Lim1] (Figure 3C); FoxA1, Lefty1, Cer1 (Figure S2J; data not shown), mesoderm (T, Mesp1, Nkx2.5; Figure 3D), and endoderm (FoxA2, Sox17, Hex1; Figure 3D) were upregulated. Marker genes specific for the pregastrula epiblast (Fgf5, Otx2, and Nanog) were also significantly upregulated after Tbx3 induction (Figure S2K). In contrast, ectoderm (Pax6, Fgf4, Noggin, Gbx2) and trophectoderm-specific genes (Cdx2, Hand1, Wnt6, Fgf2) were reduced (Figures 3E, S2L, and S2M).

Small-hairpin-RNA (shRNA)-mediated knockdown of TBX3 overexpression on marker genes of all three germ layers (Figures 3G and S2O) at the expense of trophectoderm class genes, whereas no specific regulation of ectoderm class genes was observed (Figure 3H). Also, principle component analysis and hierarchical clustering positioned the transcriptome of Tbx3-induced samples within the three germ layers toward DE (Figures 3I and 3J). In summary, we show that Tbx3 induces a subset of marker genes labeling mesoderm and endoderm and serves as a regulator of the mesendodermal transcriptional program.

(B) IF of E3.5 blastocyst for OCT4 (red) and Tbx3 (green). Scale bars, 20 μm. Nuclei in DAPI (blue).
(C–E) Tbx3 expression in postimplantation mouse embryos. (C) Tbx3 mRNA in situ hybridization at E6.25 (prestreak) and E7.5 (midstreak). exVE, extraembryonic visceral endoderm. Arrows point to Tbx3-positive cells within the proximoposterior epiblast at E6.25 and mesoderm at E7.5. (D) Tbx3 mRNA in whole mounts of wild-type embryos at E7.75 and E8.25. (E) Z-section of an E6.5 Tbx3-GFP reporter embryo. Nuclei in DAPI (blue). Scale bars, 100 μm. (F and G) tbx3 expression in Xenopus. (F) Temporal expression of tbx3 during early Xenopus embryogenesis. tbx3 transcripts are maternally supplied. Zygotic expression starts at stage 10 and increases during gastrulation. gapdh as loading control. Negative control represents a −RT reaction with gapdh. (G) Spatial expression of tbx3 at stages indicated during Xenopus development. tbx3 is detected in the invaginating mesoderm during gastrulation. White arrowheads, tbx3-positive ectodermal cells; red arrowheads, tbx3 expression in the developing cardiac region. m, mesoderm; cm, cardiac mesoderm. Scale bars, 0.5 mm. See also Figure S5.
**TBX3 Directs Cell Fate toward Mesendoderm-Derived Tissues**

We also found TBX3 to be expressed in early cardiac progenitors during mouse and *Xenopus* development (Figures 1D and 1G). We therefore assessed whether late mesendoderm-derived tissues were likewise induced upon TBX3 overexpression in prolonged ESC differentiation cultures. At late time points of differentiation, EBs overexpressing TBX3 showed an increase in pancreatic differentiation indicated by *Hnf1α* and *Pdx1* RNA levels (Figure 4A). IF shows few scattered PDX1-positive cells under control conditions, whereas large clusters of pancreatic progenitors are detected in Dox-induced cultures (Figure 4B). Similarly, increased hepatic differentiation is shown by elevated *Hnf4α* and *Albumin* (*Alb*) RNA levels (Figures 4C and 4D). For mesoderm derivatives, TBX3 induction during the first 2 days of EB differentiation significantly enhanced the generation of beating cardiomyocytes (Movies S1 and S2; Figures 4E–4H), accompanied by increased mRNA levels of early cardiac-specific transcription factors *Mesp1* and *Nkx2.5* (see Figure 3D). Markers for the first and second heart field were induced to a similar extent (Figure 4E). Interestingly, cardiac induction was most efficient when TBX3 was induced early (days 0–2 and 0–4) during differentiation, whereas TBX3 induction following day 4 did not further promote cardiac differentiation. Atrial and ventricular markers were expressed at comparable levels (Figure 4F). We further confirmed cardiac induction by protein expression of *α-ACTININ* (Figure 4G) and assessment of beating clusters as well as beating intensity under the respective culture conditions (Figure 4H). Vascular markers (*CD31*, *VEGF-A*, *KDR*) were additionally increased, indicating that TBX3 acts early during the generation of the common cardiovascular progenitor (Figure 4I). Taken together, our data establish that expression of TBX3 early in differentiation promotes enhanced formation of definitive endoderm and mesoderm as well as respective later stage derivatives, namely pancreas, liver, and heart.

**TBX3 Acts via a NODAL Signaling Dependent, Non-Cell-Autonomous Mechanism to Direct Mesendodermal Cell Fate**

Secreted signaling molecules such as WNT3a, BMP4, and NODAL are required for lineage commitment during gastrulation onset (Zorn and Wells, 2009). Thus, we further investigated our microarray data to identify an impact of TBX3 on any of these three factors. Interestingly, mRNA expression of *Nodal* and its target genes were significantly increased in TBX3-overexpressing cells, whereas *Bmp4* was significantly downregulated (Figures 5A and 5B). Because NODAL signaling acts as one of the key pathways required for induction of mesendoderm both in vivo and in vitro (Arnold and Robertson, 2009), we investigated the relationship between TBX3 and NODAL signaling.

*Nodal* expression levels were significantly upregulated in TBX3-overexpressing cells via qPCR (Figure 5C). Moreover, phosphorylated SMAD2/3 levels, the major transducer of activated NODAL signaling, were significantly increased in TBX3-overexpressing cells (Figures 5E and 5F). We analyzed direct (responding to SMAD2/3 activation in the absence of protein synthesis) and indirect SMAD2/3 target genes (presence of protein synthesis) (Guzman-Ayala et al., 2009) and found both to be enriched in TBX3-overexpressing samples (Figures 5G and 5H). Blocking NODAL/SMAD2 signaling (ALK4/5/7 inhibitor) significantly attenuated the TBX3-inductive effects as shown by reduced expression of *Eomes*, *Nodal*, *Hex1*, and *FoxA2* (Figure 5I). To address binding of TBX3 on the *Nodal* promoter, we performed ChIP using TBX3-specific antibodies. We found a significant TBX3 enrichment on a highly conserved T-box binding element within the promoter region of the *Nodal* gene (Figure 5D; Figures S4A and S4B).

Due to the inductive impact of TBX3 on NODAL expression, we hypothesized that TBX3 affects cell specification via enhancing paracrine NODAL signaling (non-cell-autonomous mechanisms). We combined fluorophore-labeled iTBX3 cells (expressing the fluorophore tdTomato...
TBX3 Drives Mesendoderm
after lentiviral transduction) and unlabeled control cells in a chimeric EB culture system. Both cell lines were mixed at the indicated ratios, and TBX3 expression was induced by Dox. Nontreated cells served as controls (Figure S3A). Labeled and unlabeled cells were subsequently separated by FACS, and marker genes were quantified. We found significantly increased expression of early mesendoderm markers, such as Lefty1 and Eomes (Figure S3B) in TBX3-overexpressing cells but interestingly also in cells not overexpressing TBX3. Notably, these effects increase with rising numbers of TBX3-overexpressing cells in the mixed EBs. Similar effects were observed when conditioned medium (CM) from iTBX3 cells was applied to wild-type EBs (Figure 6A), showing a marked increase in the mRNA levels of Eomes, T, FoxA2, and Mesp1 (Figure 6B) and increased protein levels of EOMES, T, and SOX17 (Figures 6C–6E). Of note, non-CM-treated wild-type cells behaved similar to –Dox CM (Figures 6C–6E). In line, the non-cell-autonomous effect of iTBX3-CM was also observed for late mesendoderm derivatives such as cardiac cells or pancreatic progenitors (Figures S3C–S3E).

NODAL regulates its own expression but also induces negative feedback loops in which e.g., LEFTIES inhibit NODAL activity (Arnold and Robertson, 2009). To verify this physiological loop in CM experiments (Figure 6A), we identified specific upregulation of Nodal itself, as well as Lefty1 and Lefty2 in samples treated with CM generated from TBX3-overexpressing cells (Figure 6F). This strongly points to increased levels of secreted NODAL protein upon TBX3 expression. To elucidate the requirement of secreted NODAL within our model, we used a NODAL antibody to block secreted NODAL protein in iTBX3 cultures. We observed a substantial reduction in the TBX3-mediated expression of several mesendodermal genes including Eomes, T, Sox17, and FoxA2 as well as Nodal itself (Figure 6G; Figure S3F). Overall, these experiments provide evidence that the lineage promoting effects of TBX3 can be attributed in part to paracrine effects, whereby TBX3 alters expression of key developmental factors, in particular, secreted NODAL protein.

**TBX3 Directly Regulates Key Mesendoderm Determinants**

Provided TBX3 actions could be only partially explained via paracrine signaling mechanisms, we decided to further investigate cell-autonomous effects of TBX3 in mediating cell lineage specification. We measured direct transcriptional activity of TBX3 on the regulation of key factors for mesendoderm (EOMES), mesoderm (T), and DE (SOX17) development. Overexpression of TBX3 led to the activation of EOMES, T, and SOX17 reporter constructs (Figure 7A). To confirm a direct binding of TBX3 to the corresponding promoters in differentiating ESCs, we performed ChIP using TBX3-specific antibodies. We found significant TBX3 enrichment on highly conserved T-box binding elements (Figure S4A), specifically within the promoter regions of Eomes, T, and Sox17 (Figures 7B–7D; Figures S4C–S4E) with no enrichment for Fg2 (Figure 7E). Thus, our data suggest that TBX3 can directly or indirectly activate a mesendodermal lineage specification program via transcriptional regulation of specific mesoderm and endoderm transcription factors, and partly through activation of a NODAL/SMAD2 signaling cascade.

OCT4 has been shown to bias differentiation toward mesendoderm in ESCs and is known to activate the TBX3 promoter (Thomson et al., 2011). We asked whether TBX3 also regulates OCT4 levels during ESC differentiation. Indeed, OCT4 mRNA and protein expression levels were significantly elevated at day 2 and day 4 of differentiation following TBX3 induction (Figures 7F and 7G). Regulation of OCT4 by TBX3 was further corroborated using reporter assays, whereby TBX3 overexpression significantly enhanced OCT4 reporter activity (Figure 7H). The direct occupancy of TBX3 on the OCT4 promoter and vice versa has been shown previously (Figure S1; Thomson et al., 2011).

**Figure 3. TBX3 Promotes Mesoderm and Endoderm**

(A) Illustration of the differentiation protocol. EBs were generated over time from pluripotent ESCs at day 0 (orange) and markers from all three germ layers (ectoderm, blue; mesoderm, yellow; endoderm, green) were generated in –Dox samples, whereas mesoderm and endoderm were upregulated in the +Dox condition. (B) Whole-mount IF in iTBX3 EBs (day 4) for Eomes (left panel), T (middle panel), and Sox17 (right panel) (all in red) upon TBX3 induction (+Dox). All scale bars are 20 μm. Nuclei in DAPI (blue).

(C–E) qPCR analysis showing expression of marker genes at day 4 of differentiation: (C) mesendoderm, (D) mesoderm and endoderm, (E) trophoblast and ectoderm upon TBX3 induction via doxycycline (+Dox). n = 3 for all experiments in biological replicates.

(F) TBX3 shRNA shows reduced levels of Tbx3, T, and Eomes expression levels already at EB day 1 after a 4 day feeder-free ESC condition. n = 2 for all experiments in biological replicates.

(G and H) Genome-wide transcriptional profiling followed by gene set enrichment analysis highlights that TBX3 selectively enriches mesoderm and endoderm (G) at the expense of trophoblast genes (H).

(I) Principal component analysis (PCA) positions TBX3-induced cells (TBX3_d4+dox) close to endoderm.

(J) Hierarchical clustering of TBX3-expressing EBs overlaps with published transcriptomes from DE (Christodoulou et al., 2011).

*p < 0.05; **p < 0.01; ***p < 0.001. See also Figure S2.
Finally, we overexpressed OCT4 alone or in combination with TBX3 and investigated effects on promoter activity of EOMES and SOX17 luciferase reporter constructs. In line with previous reports (Stefanovic et al., 2009), elevated OCT4 levels in ESCs enhanced the promoter activity of *Eomes* and *Sox17*. Coexpression of TBX3 and OCT4 significantly increased the luciferase activity of both promoters (Figure 7I).

**TBX3 Is Entwined in Complex Regulatory Mechanisms with Closely Related T-Box Family Members during Early Development**

Next, we aimed to investigate the role of TBX3 in *Xenopus laevis* as a vertebrate in vivo system performing morpholino-oligonucleotide (MO)-mediated loss-of-function experiments (Figures S5A and S5B). Cardiac development in *Xenopus* has recently been shown to follow a similar mechanism than in higher vertebrates (Gessert and Kühl, 2009). Injected injection of *tbx3* MO into the presumptive heart region of *Xenopus* embryos resulted in cardiac malformations, edema formation, and a reduced heart beating. Thus, TBX3 loss of function in the *Xenopus* heart faithfully recapitulates the mouse phenotype (Figures S5C–S5E).

To test the requirement of TBX3 during early lineage commitment of mesendoderm derivatives, we injected *tbx3* MO into the marginal region of the two dorsal blastomeres at the 4-cell stage. Subsequent marker gene analysis at stage 10 and 10.5 revealed downregulation of *t* and *gsc* expression, respectively (Figures S7A and S7B). This indicates requirements of TBX3 for proper mesendodermal development in *Xenopus* embryos, thereby further substantiating our in vitro findings in ESCs.

Given the central role of TBX3 in mesendoderm lineage specification in vitro, we had anticipated a more global response to the loss of TBX3 in vivo, with greater changes in the expression of mesendodermal genes. We also could not detect an obvious gastrulation phenotype upon gross inspection of embryos. Furthermore, the lack of an early gastrulation phenotype in TBX3-deficient mice (Davenport et al., 2003) prompted us to hypothesize a scenario where potential compensatory mechanisms may be present. To identify potential compensatory factors, we reanalyzed several T-box factors in terms of their evolutionary proximity (Figure S6A). The analog TBX2 has previously been implicated to show functional redundancy with TBX3 during atrioventricular canal formation (Singh et al., 2012). Accordingly, TBX3 shows the highest relationship with TBX2 and to a lesser degree with TBX5 or TBX4. From a developmental point of view, TBX6 shows the highest overlap with TBX3 in expression in vitro (Pereira et al., 2011) and in vivo (Hadjantonakis et al., 2008), albeit evolutionary relationship with TBX3 is not that pronounced (Figure S6A). Based on this, we chose TBX2 and TBX6 for further analysis. In line with our hypothesis, we initially analyzed the expression of TBX2 and TBX6 in response to TBX3 overexpression in iTBX3-ESCs. Upon TBX3 induction, expression levels of TBX2 were almost abolished as shown by mRNA and protein levels (Figure S7C). Similar reductions on protein levels, but not as pronounced on mRNA levels were detected for TBX6 (Figure S7C).

We finally performed loss-of-function analysis in *Xenopus* using different combinations of MOs specifically targeting *tbx2* (Cho et al., 2011), *tbx3*, and *tbx6* (Tazumi et al., 2008) and analyzed gastrulation in more detail. Loss of TBX3 or TBX2 individually resulted in a delay of gastrulation movements as indicated by a moderately opened blastopore at stage 13 when control MO-injected embryos already closed the blastopore (Figure S7D). During subsequent development, these moderately opened blastopores closed, explaining the lack of any gastrulation phenotype at later stages. Only in some cases, the phenotype was severe and affected embryos did not recover. However, upon coinjection of *tbx2* and *tbx3* MO almost all embryos revealed a gastrulation phenotype. Taken together, these data indicate the functional redundancy of TBX3 and TBX2. Interestingly, loss of TBX6 resulted neither in delayed gastrulation movements, nor in a further increase in the percentage of phenotype or the severity of the phenotype (data not shown), indicating the specificity of the effects observed upon loss of TBX2 or TBX3. In summary, these data indicate the requirement of T-box transcription factors for proper gastrulation movements and that some T-box factors, e.g., TBX2 and TBX3 act in a redundant fashion in this context.

In summary, TBX3 is not only required to maintain pluripotency, but also to promote lineage determination via cell-autonomous regulation of key lineage determining transcription factors, at least in part additive to OCT4. Moreover, TBX3 acts non-cell-autonomously via direct activation of NODAL signaling, which, in turn, impacts...
on the specification of DE and anterior mesoderm derivatives (Figure S6B).

**DISCUSSION**

TBX3 represents a unique member of the T-box family of transcription factors, due to its role in maintaining ground state pluripotency of murine ESCs. Recently, subsets of pluripotency-maintaining factors, namely, NANOG, OCT4, KLF5, and TBX3, have been shown to adopt new roles during lineage specification and have been globally defined as “mesendoderm class embryonic stem cell genes.” However, detailed underlying insights how these transcription factors orchestrate cell-fate decisions, remain largely elusive. In the current study, we aimed to dissect the molecular function of TBX3 during early embryonic development.

The preimplantation expression pattern of TBX3 matches well with other mesendodermal pluripotency factors, such as OCT4 or KLF5 (Thomson et al., 2011), which together colocalize in the ICM of the late-stage blastocyst. Subsequently, TBX3 expression becomes rapidly restricted to the posterior proximal pole of the epiblast before a visible primitive streak has formed. Similarly, we found tbx3 expression at the dorsal blastopore lip in *Xenopus*. TBX3 expression within the posterior side of the gastrulating embryo correlates well with the expression of NODAL, BRACHYURY, EOMES, and OCT4 (Morris et al., 2012). Mid-streak embryos (E7.5) display robust expression of TBX3 in the proximal mesoderm. In ESCs, endogenous TBX3 expression coincides with crucial mesendoderm markers (EOMES, T, SOX17) and directly activates their promoters to drive mesendoderm. Further, TBX3 acts in concert with the histone demethylase JMJD3 to allow the activation of the NODAL signaling pathways, both SMAD2-dependent and -independent, are crucially involved in defining the polarity of the pregastrulation embryo (Brennan et al., 2001). This requires interactions between the epiblast and the two extraembryonic tissues. The non-cell-autonomous TBX3 function to induce mesendoderm via enhanced NODAL secretion possibly involves the TBX3-expressing cells of the ExVE. We identify a central role for NODAL as an immediate target of TBX3 actions and demonstrate that blockade of the NODAL/SMAD2 pathway significantly retains TBX3-inductive effects. Moreover, extensive dispersion between the VE and DE leading to epiblast-derived organs containing extraembryonic elements is present (Kwon et al., 2008).

TBX3-null mice die around developmental stage E12.5 due to both cardiac and yolk sac defects. These mice exhibit a smaller sinoatrial node and die from AV blockage and ventricular bradycardia (Frank et al., 2012). Our data on TBX3 loss of function in the *Xenopus* heart faithfully recapitulate this phenotype. To our knowledge, a precise analysis of TBX3 mutant mouse embryos at early gastrulation stages has not yet been performed, which could potentially identify previously unrecognized spatiotemporal
abnormalities of primitive streak formation. In fact, monogenic knockouts of certain genes result in moderate or no gastrulation defects, whereas combined ablation strategies result in more severe gastrulation defects (Solloway and Robertson, 1999; Song et al., 1999).

We found that loss of either TBX2 or TBX3 but not TBX6 results in a delay of gastrulation in *Xenopus*. We therefore propose that T-box factors act in a redundant but also specific manner to allow ongoing gastrulation. Notably, T-box genes are known to operate in a complex network to regulate region- and species-specific gene expression and developmental fate (Goering et al., 2003). Accordingly, T-box factors have now been classified into three subtypes, namely, additive, competitive, and combinatorial (Goering et al., 2003). They show overlapping expression patterns in numerous tissues, while exhibiting distinct spatiotemporal expression and functions in other compartments. This has several implications for dissecting the individual contribution of a single T-box gene within specific processes and highlights that any genetic manipulation of T-box genes such as gain- or loss-of-function analysis must be carefully taken to be context dependent.

Specifically, functional redundancy between the paralogous TBX2 and TBX3 has previously been identified during atrioventricular canal, mammalian secondary palate formation (Singh et al., 2012) as well as during *Xenopus* eye development (Takabatake et al., 2002). In this respect, TBX2/TBX3 double-knockout mice display greater overall embryonic growth defects (lethal as early as E9.5) compared to single-knockout mice (lethal at E12.5) (Mesbah et al., 2012). Tbx6 knockout mice exhibit reduced NODAL levels in early development stages (Carreira et al., 1998). This molecular abnormality would appear to somewhat overlap with TBX3 functions. However, during *Xenopus* gastrulation we could not detect any functional redundancy between TBX3 and TBX6.

Altogether, our results describe facets of TBX3 actions within early cell-fate determination. We demonstrate that TBX3 is dynamically expressed during specification of the mesendoderm lineages in differentiating ESCs in vitro and at specific sites during gastrulation onset in developing mouse and *Xenopus* embryos. We depict that TBX3 exerts dual cell-autonomous and non-cell-autonomous effects by directly activating core mesendoderm lineage specification factors and influencing NODAL/SMAD2 signaling, respectively. Moreover, we show that TBX3 acts together with OCT4 to promote mesendoderm specification. Finally, we display a functional redundancy between TBX3 and the closely related family member TBX2, pointing to a complex compensatory mechanisms. We propose a model in which TBX3 orchestrates a complex network of downstream effects within the T-box family members to direct pluripotent cells toward a mesendodermal fate. A more comprehensive understanding of the role of pluripotency factors in subsequent lineage commitment may provide important insights about tightly regulated developmental processes such as gastrulation.

**EXPERIMENTAL PROCEDURES**

Full methods accompany this paper in the Supplemental Experimental Procedures.

**Generation of iTBX3 ESCs**

One day before the nucleofection procedure, A2lox.cre cells were exposed to 1 μg/ml of doxycycline to induce Cre recombination. ESCs were nucleofected using the Nucleofector Technology (Lonza) according to the manufacturer's procedures. The Nucleofector Kit for Mouse Embryonic Stem Cells was used according to manufacturer's instructions and 10 μg of DNA (TBX3-p2lox vector) per 5 million parental A2lox.cre ESCs. Nucleofected cells were plated on neomycin-resistant, Mitomycin-C (Sigma) inactivated MEFs and selected 2 days after nucleofection with neomycin (400 μg/ml).

**Gene Expression Microarray Analysis**

The Agilent DNA microarray data were preprocessed and normalized by the limma package (Smyth, 2004). The Affymetrix DNA microarray data (GSE27087) were normalized by RMA algorithm implemented in Affymetrix power tools. The investigated gene sets were collected from Mouse Genome Informatics database (Finger et al., 2011) and literature.

**Figure 6. TBX3 Acts via Non-Cell-Autonomous Mechanisms to Direct Cell Fate**

(A) Illustration of the treatment strategy to generate conditioned medium (CM) from iTBX3 cells and subsequent generation of EBs. (B) qPCR for *Eomes, T, Foxa2*, and *Mesp1* in differentiating wild-type ESCs cultured in either non-CM or +Dox-CM days 0–4 from iTBX3 cells on day 4.

(C–E) IF of differentiating wild-type ESCs cultured in either non-CM or CM from iTBX3 cells (either ±Dox) on day 4; wild-type cells treated with +Dox-CM days 0–4 of iTBX3 cells show a higher protein expression of the depicted markers (C) EOMES, (D) T, and (E) SOX17. All scale bars are 20 μm. Nuclei in DAPI (blue). n = 3 in biological replicates.

(F) qPCR analysis for *Nodal, Lefty1*, and *Lefty2* mRNA levels in either non-CM or +Dox-CM days 0–4 from iTBX3 cells on day 4.

(G) *Nodal, Eomes, T*, and *Sox17* mRNA levels in untreated (−Dox) and Dox-treated (+Dox) iTBX3 cells at day 4 of differentiation in ± 3 μM NODAL-blocking antibody, (n = 2) independent experiments, each in biological triplicates.

*p < 0.05; **p < 0.01; ***p < 0.001. See also Figure S3.
Figure 7. TBX3 Directly Regulates Key Factors for Mesendoderm Induction

(A) Luciferase reporter of EOMES (left bar), T (middle bar), and SOX17 (right bar) promoters were cotransfected in HEK cells with a TBX3-expression plasmid and assessed for luciferase activity, (n = 3) biological replicates. Similar data with iTBX3 ESCs (data not shown).

(B–E) qPCR quantification of DNA fragments enrichment by ChIP using TBX3-specific antibody relative to control isotype for (B) Eomes, (C) T, (D) Sox17, and (E) Fgf2. n = 3 for biological replicates.

(legend continued on next page)
Quantitative One-Step Real-Time RT-PCR
Quantitative one-step real-time RT-PCR analysis was implemented as previously described (Kleger et al., 2007, 2010; Liebau et al., 2007).

Chromatin Immunoprecipitation
Approximately, $2 \times 10^6$ differentiated ESCs were crosslinked by adding formaldehyde directly to the medium to a final concentration of 1% (w/v) at room temperature. The ChIP Assay kit (ChIP-IT Express Enzymatic, Active Motif, #53009) was used according to the manufacturer’s instructions. Details are in the Supplemental Information.

Statistical Analysis
If not stated otherwise, SEMs are indicated by error bars. For comparisons of two groups, levels of significance were calculated with the unpaired Student’s t test if not indicated otherwise. *p < 0.05; **p < 0.01; ***p < 0.001. Significances were calculated with Prism5 (GraphPad).

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, seven figures, and two movies and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2013.08.002.

ACKNOWLEDGMENTS
The authors thank R. Köhntop, S. Fischer, and S. Seltenheim for excellent technical assistance. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, KL 2544/1-1, SL BO1718/4-1, MZ Ze432/5-2), German Foundation for Heart Research (F/34/11; to A.K. and S.L.), Boehringer-Ingelheim BIU (N5 to S.L.; C1 to A.K.), Helmholtz Gesellschaft (VI-VI-S10 to T.M.B. and S.L.), Else-Kröner-Fresenius-Stiftung (2011_A200; to A.K. and S.L.), British Heart Foundation (BHF RG 04/003951; to A.K.), Bundesministerium für Bildung und Forschung (BMBF, MND-Net to S.L. and T.M.B., Systar to A.K.), Helmholtz Gesellschaft (VH-VI-510 to T.M.B. and S.L.), German Foundation for Heart Research (F/34/11; to T.M.B. and S.L.), German Research Foundation (DFG, KL 2544/1-1, SL BO1718/4-1, MO 2366/26-1), Else-Kröner-Fresenius-Stiftung (2011_A200; to A.K. and S.L.), and German Research Foundation (KL 2544/1-1, SL BO1718/4-1, MO 2366/26-1). Excellent technical assistance. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, KL 2544/1-1, SL BO1718/4-1, MZ Ze432/5-2), German Foundation for Heart Research (F/34/11; to A.K. and S.L.), Boehringer-Ingelheim BIU (N5 to S.L.; C1 to A.K.), Helmholtz Gesellschaft (VI-VI-S10 to T.M.B. and S.L.), Else-Kröner-Fresenius-Stiftung (2011_A200; to A.K. and S.L.), BMBF (MND-Net to S.L. and T.M.B.; Systar to A.K.) and the Baden-Württemberg Stiftung (BioEqual, C.A., Gadue, P., Spira, A., Gouon-Evans, V., Murphy, G.J., Moshtalovsly, G., and Kotton, D.N. (2011). Mouse ES and iPS cells can form similar definitive endoderm despite differences in imprinted genes. J. Clin. Invest. 121, 2313–2325.

Received: June 18, 2013
Revised: August 6, 2013
Accepted: August 7, 2013
Published: August 29, 2013

REFERENCES


(F and G) TBX3 induces OCT4 expression: (F) mRNA level of OCT4 upon TBX3 induction (±Dox) at indicated time points. (G) Protein level of OCT4 whole-mount stained EBs upon TBX3 induction (±Dox) at day 4. All scale bars 20 µm. Nuclei in DAPI (blue). n = 3 for biological replicates.

(H) Schematic outline of OCT4 reporter assays (left panel). Luciferase reporter constructs with either proximal or distal OCT4 enhancer elements cotransfected into HEK cells together with a TBX3-expression plasmid and assayed for luciferase activity. n = 3 for biological replicates. Similar observations were made when iTbx3 ESCs were used (data not shown).

(I) HEK cells were cotransfected with the respective luciferase reporter constructs in the presence or absence of OCT4- and TBX3-expression plasmids (schematic outline, left panel) and assayed for promoter activity of Sox17 (middle panel) and Eomes (right panel). Similar observations were made when iTbx3 ESCs were used (data not shown). n = 3 for biological replicates.

*p < 0.05; **p < 0.01; ***p < 0.001. See also Figures S4 and S7.


Stem Cell Reports, Volume 1

Supplemental Information

TBX3 Directs Cell-Fate Decision toward Mesendoderm

Supplemental Information

Inventory
Supplemental Figures 1-7
Supplemental Figure Legends
Supplemental Experimental Procedures
Supplemental References
**Supplemental Figure 1 (related to Figure 2). TBX3 binds lineage determining genes in the pluripotent state.** Solexa ChIP-sequencing data sets, which have been generated elsewhere (Han et al., 2010), to uncover the direct regulatory targets of TBX3 during reprogramming, were reanalyzed for TBX3 binding of certain pluripotency genes (dark grey columns; *Klf5, Klf4, Oct4*) and lineage determining genes (light grey; *Sox9, Nkx6.1, Eomes*). X-axis shows fold enrichment of certain genetic regions of the respective genes upon chromatin immunoprecipitation with a TBX3-binding antibody.
Supplemental Figure 2 (related to Figure 3). Generation of a targeted, inducible TBX3-expression allele in ESCs. (A) Scheme of cloning strategy to generate a temporally regulated and dose-dependent Dox-inducible TBX3 ESC line. (B) High expression of Tbx3 mRNA in Dox-treated iTBX3 ESCs as assessed by qPCR after 24 and 48 hours. (C,D) TBX3 protein levels are increased in response to Dox-treatment after 24 hours in iTBX3 ESCs and EBs respectively. All experiments (n=3) in biological replicates. TBX3 tightly regulates certain lineages. (E,F) IF of T and SOX17 demonstrate TBX3-induction enhances differentiation towards mesendodermal derivatives respectively. All scale bars are 20 μm. (G,H) Respective quantifications of either (E) or (F). (I) WB analysis of day4 and day6 EBs, confirms TBX3 (+Dox) induces SOX17-expression levels more compared to control (-Dox) cells. All experiments (n=3) in biological replicates. (J-M) Microarray expression data depicted as log2 fold for a series of lineage specific genes in day4 -Dox or +Dox samples. TBX3-induced ESCs display upregulation of (J) early mesendodermal marker genes and (K) epiblast–associated genes while (L) neuroectodermal and (M) trophoblast genes are downregulated upon TBX3-induction. (N) qPCR analysis of certain marker genes labeling primitive endoderm. (O) Gene set enrichment plot highlights that TBX3-induction favors primitive endoderm formation. p-values are represented as <0.001, ***; <0.01, **; <0.05, *. 
Supplemental Figure 3 (related to Figure 6): (A,B) Wild-type and td-tomato-labeled iTBX3 cells are mixed at the indicated ratios and chimeric EBs are formed. After differentiation (-/+Dox) for 6 days, single cells from chimeric EBs are FACS sorted and analyzed by qPCR as shown in (B). All experiments (n=3) in biological replicates. (B) qPCR analysis for Lhx1 (upper panel) and Eomes (lower panel) FACS-purified as outlined in (A) for the indicated populations. High expression levels of Lhx1 and Eomes in both TBX3 and wild-type cells are associated with higher TBX3 expression in the origin EBs. All experiments (n=3) in biological replicates. (C) The number of beating cardiomyocyte clusters per well are counted on day14 in wild-type ESCs treated either with non-CM, with -Dox CM or with +Dox-CM from iTBX3 cells (+Dox CM day0-2 or day0-4). Higher levels of (D) the cardiac marker gene Myh6 and (E) the pancreatic progenitor marker gene Pdx1 at day14 are associated with higher TBX3 levels as assessed by qPCR in wild-type ESCs treated either with non-CM, with -Dox CM, or with +Dox-CM from iTBX3 cells (+Dox CM day0-2 or day0-4). (F) qPCR analysis for FoxA2 levels in untreated and Dox-treated iTBX3 cells at day4 of differentiation in the presence or absence of Nodal-blocking antibody at a final concentration of 3 μM (n=2) in biological replicates. p-values are represented as <0.001,***; <0.01, **; <0.05, *. 
Supplemental Figure 4 (related to Figures 5 and 7): Highly conserved promoter regions and ChIP primer localization. (A) Consensus T-box binding site generated by MEME (http://meme.nbcr.net/meme/cgi-bin/meme.cgi). Fw: forward sequence; rc: reverse complimentary sequence. (B-E) Alignment of mouse genomic loci of *Nodal* (B), *Eomes* (C), *T* (D) and *Sox17* (E) to the respective orthologous using rVista browser. The putative Tbx3 binding sites were detected using position weight matrix (PWM) of the Tbx-motif generated in (A). Two conserved binding sites for each gene promoter are shown in the respective table. The genomic location for the mouse gene is given together with the respective binding score and orientation of the respective motif (fw, rc). Angular black arrow indicates the TSS (transcriptional start site) and blue dotted box marks the genomic region containing the two depicted conserved motifs and the ChIP-PCR amplicon. Black dotted box zooms into the blue dotted box and indicates ChIP-PCR amplicon.
A
*tbx3* MO binding site

5'-A AAG TGA ATG AAT TTA CCC ATG AGA-3'
red = ATG start codon

B
*tbx3* MO-gfp+  
control MO  
tbx3 MO

C
control  
tbx3 MO

D

E
Supplemental Figure 5 (related to Figure 7). Testing the specificity of tbx3 morpholino oligonucleotide. (A) *Xenopus* tbx3 morpholino oligonucleotide (MO) binding site. The start codon is highlighted in red. (B) The tbx3 MO-binding site is cloned in front of and in frame with *gfp* and injected bilaterally as RNA together with tbx3 or control MO into *Xenopus* embryos at 2-cell stage. GFP fluorescence was monitored at stage 13. Coinjection of tbx3 MO-*gfp* RNA together with the control MO has no effect on GFP fluorescence whereas the tbx3 MO efficiently blocks translation of GFP indicating the specificity of the MO used. Scale bars are 0.5 mm. TBX3 loss of function phenotype in *Xenopus* heart development. (C) Loss of TBX3 by bilateral injection of 40 ng tbx3 MO into 84-cell stage *Xenopus* embryos results in cardiac malformation and cardiac edema at stages 42 and 45, respectively. a, atrium; v, ventricle; OFT, outflow tract. Scale bars are: lateral 1 mm, ventral 0.2 mm, isolated hearts 0.2 mm. (D) Quantitative presentation of data shown in (C). (E) Loss of TBX3 results in bradycardia at stage 42. n, number of independent experiments, N, number of embryos examined. Error bars indicate standard errors of the means (SEM). p-values are calculated by a nonparametric Mann-Whitney rank sum test. * p<0.05, ** p <0.01, **** p<0.0001.
Supplemental Figure 6 (related to Figures 1–7). T-box factors and TBX3 function hypothesis. (A) Phylogenetic relationship of T-box factors. Numbers on each branch indicate the distance to the branch node (round dot). Factors in the same branch are closely related. (B) Schematic model of proposed TBX3 action. Schematic model depicting the mechanism how TBX3 may regulate early lineage commitment and mesendodermal specification. Thus, TBX3 cell-autonomously activates T, EOMES, SOX17 and NODAL signaling components thereby activating non-cell autonomously NODAL-SMAD2/3 target genes to favor the mesendoderm/endoderm specification. The latter process possibly involves the TBX3-expressing cells of the extraembryonic.
Supplemental Figure 7 (related to Figure 4). TBX3 is functionally related to other T-box family members in directing early cell fate choices.

(A-B) Bilateral injection of 40 ng tbx3 MO leads to reduced t (A) and gsc (B) expression (black arrowheads). Uninjected (wild type, WT) and control MO injected embryos revealed normal marker gene expression. Vegetal and lateral views and sagittal sections of *Xenopus* embryos at stages 10.5 (t expression) or 10 (gsc expression) are shown as indicated. Scale bars are 0.5 mm. (C) TBX3 overexpression abolishes TBX2 and TBX6 expression on mRNA and protein level upon TBX3 induction (-/+ Dox) at day 4 of differentiation. All scale bars are 20 µm. All nuclei are shown in DAPI (blue). (D) Injection of tbx2 or tbx3 MO resulted in gastrulation phenotypes. In mild cases, closure of the blastopore was delayed in comparison to control MO injected embryos. In severe cases, gastrulation was completely blocked. n = number of independent experiments. N = number of injected embryos analyzed. Error bars indicate standard error of the means (SEM). p-values were calculated by a nonparametric Mann-Whitney rank sum test. * p<0.05, ** p <0.01. Scale bars are 0.5 mm.

**Movie S1.** Live-cell imaging of spontaneously beating areas of cardiomyocytes derived from control ES cells (related to Figure 4). Cells were imaged at day 17 of differentiation.

**Movie S2.** Live-cell imaging of spontaneously beating areas of cardiomyocytes derived from TBX3 overexpressing ES cells (related to Figure 4). Induction of TBX3 was started from day 0 to day 2 of EB generation. Note that TBX3 overexpression enhances the generation of beating cardiomyocytes. Cells were imaged at day 17 of differentiation.
Supplemental Experimental Procedures.

Cell culture & ESC differentiation.

HEK293T cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM; Invitrogen, Carlsbad, CA, USA, www.invitrogen.com), with 10% (v/v) fetal calf serum (FCS; Lonza, Basel, BS, Switzerland, www.lonza.com) and 1% Penicillin/Streptomycin (Millipore, Billerica, MA, USA, www.millipore.com). Mouse embryonic fibroblast (MEF-Feeder) cell lines were grown in DMEM with 10% (v/v) FCS (PAA, Pasching, Austria, www.paa.com), 1% Penicillin/Streptomycin, 1% GlutaMax (Invitrogen; final concentration 2mM), 1% Non-Essential Amino Acids (NEAA; Invitrogen), 1% Sodium Pyruvate (Invitrogen; final concentration 1mM), 1% β-Mercaptoethanol (Millipore) and Vitamin C (VitC; 0.05mg/ml; Sigma) in a humidified incubator containing 5% CO2 at 37°C. Feeder irradiation was performed according to standard procedures. mESCs were cultured in Knockout DMEM (KO-DMEM; Invitrogen), 15% FCS (Lonza), 1% Penicillin/Streptomycin, 1% GlutaMax, 1% NEAA, 1% Sodium Pyruvate, 1% β-Mercaptoethanol and 240 U/ml leukemia inhibitory factor (LIF; Sigma-Aldrich, St.Louis, MO, USA, www.sigmaaldrich.com).

Sox 17-RFP reporter cell line was kindly provided by Douglas Melton, T-GFP/Dppa4-RFP reporter cell line was kindly provided by Hans-Jörg Fehling. In vitro differentiation of ESCs was carried out according to the standard protocol using the hanging drop method. Briefly, Iscove’s modified Dulbecco’s medium (IMDM; Invitrogen) supplemented with 10% FCS (Lonza), 1% Penicillin/Streptomycin, 1% GlutaMax, 1% NEAA and freshly prepared Monothioglycerol (Sigma; final concentration 450 μM) were used for differentiation. In short, 600 cells per 20 μl differentiation medium were placed on the lids of petri dishes filled with 10 ml Dulbecco’s Phosphate Buffered Saline (DPBS; Invitrogen) and were cultivated for 2 days in hanging drops. In the following, embryoid bodies (EBs) were rinsed into non-adherent bacterial dishes and were cultivated for another two days. On day 4 EBs (n = 11) were plated on (0.1%) gelatin-coated 6-well dishes or cover slips for RNA or immunofluorescence analysis, respectively and assayed at specific time points as described in figure legends. Inhibitors: SB-421542 (Sigma-Aldrich, St.Louis, MO, USA, S4317, www.sigmaaldrich.com) was dissolved in DMSO and added to cell culture medium where indicated at a final concentration of 10µM. Anti-NODAL antibody (Santa Cruz, Santa Cruz, CA, USA, sc- 28913, www.scbt.com) was applied to the differentiating cells at a final concentration of 3µM.
RNA interference. TBX3 RNA interference was performed using a shRNA-expressing pSuperpuro (Oligoengine) construct as previously described in (Kartikasari et al., 2013). Briefly, mESCs were plated in feeder-free conditions prior to transfection using Lipofectamine 2000 (Invitrogen) according to manufacturer’s instructions, followed by Puromycin (Sigma) selection at 1 mg/ml for two days. EB differentiation was performed by culturing the cells in non-adherent conditions at a density of 104 cells/ml for a day in LIF-free supplemented KO-DMEM.

Lentiviral infection. Lentivirus generation was implemented by using a virus containing a polycistronic expression cassette encoding LV-tdTOMATO (kindly provided by Konrad Hochedlinger) in a 70 % confluent 10 cm dish containing Lenti-X 293T cells (Clontech, Mountain View, CA, USA, www.clontech.com) by cotransfection of the polycistronic vector (8 µg), the pMD2 vector (2 µg) and the psPAX2 vector (5.5 µg) (all Addgene, Cambridge, MA, USA, www.addgene.org) using 100 µl PolyFect transfection reagent (Qiagen, Hilden, NRW, Germany, www.qiagen.com). Viral supernatant was collected at 24, 48 and 72 hours after transfection, concentrated using the Lenti-X Concentrator Kit (Clontech).

Transfection of Reporter plasmids & Luciferase assays. At least three independent transfection experiments in duplicates were carried out in each case. For luciferase assays, HEK-293T cells were transfected in 24-well plates with 200 ng of indicated reporter plasmids and 50 ng of pTK-RL as internal transfection control, using Lipofectamine 2000 reagent (Invitrogen). Cells were lysed in passive lysis buffer (Promega, Madison, WI, USA, www.promega.com) 24 hours post transfection and luciferase activity was measured applying firefly and renilla assay buffer (Promega) using a luminometer (GloMax 96, Promega). Light emission from the firefly luciferase was normalized to light emission from the renilla luciferase for every transfection and represented as relative luciferase units (RLU).

Luciferase Reporter Constructs.

Brachyury-reporter. The Brachyury-promoter luciferase reporter construct (pTPwt: T-Promoter in pGL3Basic) was previously generated by cloning a 618 bp promoter fragment of the mouse Brachyury gene, corresponding to region -484 to +134 of the 5’region of Brachyury (Clements et al., 1996), into the promoter-less luciferase reporter plasmid pGL-3Basic (Arnold et al., 2000). We obtained this plasmid from R. Kemler.
**hEOMES-reporter.** A fragment of the hEOMES enhancer region (~6943 to ~6090) was previously generated by subcloning into the promoter-less luciferase reporter plasmid pGI-3Basic (Teo et al., 2011). We obtained this plasmid from L. Vallier.

**Sox17-reporter.** The luciferase reporter plasmid pSox17-5.6kb was previously constructed by amplifying a 5.6 kb fragment upstream of the translation start site from FVB/N mouse genomic DNA and was cloned into the promoter-less luciferase reporter plasmid pGI-3Basic (Lin et al., 2010). We obtained this plasmid from J. Wells.

**Oct4-reporter.** For quantifying relative Oct4-enhancer activities, the proximal and distal enhancers were previously PCR-amplified from a GOF18 plasmid (Yeom et al., 1996) and cloned into the promoter-less luciferase reporter plasmid pGl-3Basic (Greber et al., 2010).

**Immunocytochemistry.** Briefly, cells were fixed at different time points of differentiation using 4% Paraformaldehyde (PFA). Further they were treated with NH4Cl and blocked with BSA and 0.3% Triton. Primary antibodies were added: rabbit anti TBX3 (kindly provided by H.Niwa), 1:1000, 1h room temperature (RT); mouse anti OCT3/4 (Santa Cruz, Santa Cruz, CA, USA, www.scbt.com), 1:200, o.N. 4°, sc-5279; rat anti human/mouse EOMES (eBioscience, San Diego, CA, USA, www.ebioscience.com), 1:200, o.N. 4°, 14-4876-82; rabbit anti EOMES (Abcam, Cambridge, UK, www.abcam.com), 1:1000, o.N. 4°, ab23345; goat anti human Brachyury (R&D Systems, Minneapolis, MN, USA, www.rndsystems.com), 1:100, o.N. 4°, AF2085; goat anti human SOX17 (R&D Systems), 1:500, o.N. 4°, AF1924; goat anti human PDX1 (R&D Systems), 1:500, o.N. 4°, AF2419; rabbit anti mouse Albumin (FITC conjugated, Cedarlane, Burlington, ON, Canada, www.cedarlanelabs.com), 1:100, 1h RT, CLFAG3140; mouse anti α-Actinin (Sigma-Aldrich), 1:150, 1h RT, A7811; goat anti TBX6 (Santa Cruz, Santa Cruz, CA, USA www.scbt.com), 1:300, 1h RT; rabbit anti TBX2 (kindly provided by C. Goding). Finally fluorescence labeled secondary antibodies Alexa Flouer® 488 (green), Alexa Flouer® 568 (red), Alexa Flouer® 647 (magenta) (Invitrogen, all diluted 1:500) were added. Nuclei were stained with DAPI (1:20,000). Images were captured using an upright fluorescence Zeiss Axioimager Z1 microscope and analyzed using Axiovision software (Zeiss, Oberkochen, BW, Germany, www.zeiss.com). Embryos were flushed at 1.5 days post coitum (dpc) and cultured until 3.5 dpc in M16 media (Sigma, #M7292). Staining was performed as previously described (Nichols et al., 2009) with rabbit anti TBX3 (kindly provided by H.Niwa), 1:1000, o.N. 4° and mouse
anti OCT3/4, 1:200, o.N. 4°, sc-5279. All images were taken by an ApoTome fluorescence microscope (Carl Zeiss microscopy).

**Whole-mount in situ hybridization.** Whole-mount in situ hybridization (WMISH) analysis of E6.0 embryos and respective *Xenopus* embryo stages was performed according to standard protocols as previously described using probes for mouse and *Xenopus* TBX3 (Costello et al., 2011; Herrmann et al., 2011).

*Xenopus experiments.** *Xenopus laevis* embryos were staged according to standard protocols (Nieuwkoop and Faber, 1975). *Xenopus* embryos at different stages were fixed using Formaldehyde. WMISH experiments were performed following well established protocols (Gessert et al., 2007). Morpholino oligonucleotides (MOs) were purchased from GeneTools, LLC, USA. The *tbx3* MO was: 5'- tct cat ggg taa att cat tca ctt t – 3'. *tbx2* and *tbx6* MOs were used as previously published (Cho et al., 2011; Tazumi et al., 2008). For injection control, the standard control MO from GeneTools was used.

To investigate the temporal expression of *tbx3* during *Xenopus* development, total RNA of whole *Xenopus* embryos at diverse stages was isolated using the peqGOLD RNApure kit (Peqlab). cDNA was generated using random primers (Invitrogen) and Superscript II RNase H- reverse transcriptase (Invitrogen). RT-PCR was performed using the Master Amp Tm Taq PCR Core kit (Epicentre). Primer sequences were: *tbx3* _RT_L_: 5'- ata cca acc cag gtt cca cat agt g -3'; *tbx3* _RT_R_: 5'- gag ccc ttg aga gga aga tgc c -3'; gapdh _RT_L_: 5'- gcc gtg tat gtg gaa tct -3'; gapdh _RT_R_: 5'- aag ttg tct gtc gaa ttt gc -3'. Product length were: *tbx3*: 976 bp; *gapdh*: 230 bp. Annealing temperature was 55°C. To examine the spatial expression of *tbx3*, a *tbx3* cDNA probe with a length of 976 bp was cloned using the proof reading PfuUltra II DNA polymerase (Stratagene) and the following primers: *tbx3* _L_: 5'- ata cca acc cag gtt cca cat agt g -3'; *tbx3* _R_: 5'- gag ccc ttg aga gga aga tgc c -3'. The PCR product was cloned into the pSC-B vector (Stratagene). A *tbx3* antisense RNA probe was generated using NotI (NEB) and T3 RNA polymerase (Roche). To detect the spatial expression of *tbx3* during gastrulation, the specificity of the *tbx3* MO was tested by cloning the binding site in front of and in frame with GFP. *In vitro* transcribed RNA was injected together with *tbx3* or control MO at 2-cell stage and fluorescence was monitored 2 days later (**Supplemental Figure 5**). MOs were injected bilaterally into the marginal zone of each dorso-vegetal blastomere at doses indicated. To analyze marker gene expression, embryos were fixed at stage 10 (*gsc* expression) and 10.5 (*t* expression). Sagittal sections were cut using a scalpel. To analyze the heart
phenotype, tbx3 MO was bilaterally injected into 8-cell stage embryos in both dorso-vegetal blastomeres to target cardiac tissue. Correctness of injections was verified by gfp RNA co-injections. Heart morphology was analyzed at stage 42 and 45, respectively. Heart beat was counted at stage 42.

**Immunoblotting.** Immunoblotting was performed according to standard procedures. Primary antibody dilutions were: rabbit anti TBX3 (Abcam), 1:1000, o.N. 4°, ab66306; goat anti TBX3 (Santa Cruz), 1:1000, o.N. 4°, sc-17871; goat anti human SOX17 (R&D Systems), 1:500, 1h RT, AF1924; rabbit anti Phospho SMAD2 (Cell Signaling, Danvers, MA, USA, www.cellsignal.com), 1:1000, o.N. 4°, #3101; β-actin, 1:50.000, 1h RT. This was followed by incubation with secondary horse radish peroxidase–coupled antibodies diluted 1:3000, 1h RT. Detection was performed with either ECL or ECL+ kits (Thermo scientific, Waltham, MA, USA, www.thermofisher.com).

**FACS analysis.** On designated days, EBs were washed with PBS and dissociated into single cell suspension by incubation with 0.25% trypsin/EDTA (Millipore). The cells were dissolved in 5% FCS/PBS and analyzed with BD FACSaria III cell sorter. All events were gated with forward scatter (cell size) and side scatter (cell granularity) profiles.

**Bioinformatic analysis.**
The ECR browser (http://ecrbrowser.dcode.org) tool was used in the rVista genome browser (http://rvista.dcode.org) to identify the evolutionarily conserved DNA sequences within the promoters of Nodal, Eomes, Brachyury (T) and Sox17 followed by Tbx-binding site prediction analysis. All the putative binding site sequences from different promoters were retrieved and subjected for de novo consensus Tbx-binding motif generation using MEME motif generation tool (meme.ncbr.net/meme/). All genomic and transcript sequences were obtained from either UCSC browser (http://genome.ucsc.edu/) or Ensembl (http://useast.ensembl.org/index.html). The Tbx-motif was displayed as the motif logo by seqLogo package in R. To estimate the Tbx3-binding of Eomes, Nodal, T and Sox17 in the mouse genome (Mouse GRCm38/mm10 assembly), the putative Tbx-binding sites in the flanking regions were detected for 600 up- and down-stream positions around ChIP primer binding sites. The motif match was performed using matchPWM in Biostrings package in R. The minimum score for counting a match was set to 0.8 indicating 80% of chance that the predicted site is a Tbx-binding site. Afterwards the motif conservation among
human, rat and mouse genomes was assessed using Vertebrate Multiz Alignment and Conservation track in UCSC genome browser.

To perform cross-platform analysis, data from different array platforms were integrated by NCBI gene id, and further adjusted by the Combat algorithm (Johnson et al., 2007) to minimize the batch effect. Principal component analysis and hierarchical clustering were performed in R programming environment (Team, 2005). Differentially expressed genes were detected by limma t-test with criteria of p-value < 0.05 and fold change > 1.5. The p-value was adjusted by the procedure of Benjamini and Hochberg (Benjamini and Hochberg, 1995). The gene set enrichment analysis was done by customed R-GSEA script (Subramanian et al., 2005).

**ChIP detailed information.** Antibodies were applied as follows: goat anti TBX3 (Santa Cruz), o.N. 4°C, 3µg, sc- 31657; goat anti TBX3 (Santa Cruz), o.N. 4°C, 3µg, sc-17871; rabbit anti TBX3 (Invitrogen), o.N. 4°C, 3µg, #42-4800; goat IgG (Santa Cruz), o.N. 4°C, 3µg, sc-2028. TBX3 antibodies showed similar results in qPCR. The above immunoprecipitated DNA was eluted and reverse cross-linked. The purified DNA and 1% of the respective input DNA were used as templates for qPCR using Rotor Gene rtPCR Cycler (Qiagen) and SYBR Green qPCR mix (Fermentas, Burlington, ON, Canada, http://www.thermoscientificbio.com/fermentas/). Ct-values were first normalized to the respective input DNA and are represented as % of input. Appointed annealing temperatures: **Nodal** primer - 56°C, **Eomes** and **T** primer - 60°C, **Sox17** primer - 65°C. Primer information are as follows:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Sequence</th>
<th>Product length (Bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brachyury</td>
<td><strong>Fwd</strong> 5’ ccc gcc tgc tgc ccc tgc 3’&lt;br&gt;<strong>Rev</strong> 5’ agc cct gcc tcc taa act tga act 3’</td>
<td>348</td>
</tr>
<tr>
<td>Eomes</td>
<td><strong>Fwd</strong> 5’ agg gaa ttc tga tga alt aaa glg 3’&lt;br&gt;<strong>Rev</strong> 5’ ctg act tgc gat tgl tgg cag g 3’</td>
<td>425</td>
</tr>
<tr>
<td>Nodal</td>
<td><strong>Fwd</strong> 5’ ccc ccc ccc cca tcc tcc tcc ccg ctg acg cg 3’&lt;br&gt;<strong>Rev</strong> 5’ ctc cgg aga ggc cta taa cct a 3’</td>
<td>123</td>
</tr>
<tr>
<td>Sox17</td>
<td><strong>Fwd</strong> 5’ gag ggt gct gct gag tgg tgg cag g 3’&lt;br&gt;<strong>Rev</strong> 5’ cag cag cag cag tgt gtt gtt ctt cgt cgg 3’</td>
<td>223</td>
</tr>
<tr>
<td>Fgf2</td>
<td><strong>Fwd</strong> 5’ aca gac aca cag aat cag acc aac c 3’&lt;br&gt;<strong>Rev</strong> 5’ tca gta gta gta agg atg tgg 3’</td>
<td>125</td>
</tr>
</tbody>
</table>

**Quantitative one-step real-time RT-PCR and PCR detailed Information.** Briefly, one-step real-time qPCR was carried out with the Rotor Gene RT-PCR Cycler (Qiagen) using the QuantiTect SYBR Green RT-PCR kit (Qiagen). Each RNA preparation was tested for genomic DNA contamination by replacing reverse transcriptase with water. Internal standards (housekeeping gene) and samples were simultaneously amplified. RT-PCR was performed as recommended by supplier’s
For RNAi experiments, 0.5 µg of total RNA from each sample was denatured at 65°C and then reverse transcribed using Superscript II reverse transcriptase (Invitrogen) at 50°C for 1 hour. qPCR was performed in ABI7900HT (Applied Biosystems) using 1x FastSybrGreen Mix (Applied Biosystems), 500 nM of each primer and 100 ng cDNA. PCR conditions included denaturation at 95°C for 20 seconds, followed by 50 cycles of 94°C for 1 second and 60°C for 20 seconds, then continuation with dissociation stage. Primer sequences were listed in (Kartikasari et al., 2013).

Relative transcript expression is depicted as the ratio of target gene concentration to the housekeeping gene Hydroxymethylbilane synthase (Hmbs), cyclophilin or polymerase 2A (Pol2a) or for Xenopus PCR gapdh (glycerinaldehyde-3-phosphate dehydrogenase) concentration. Self-made (Biomer) and QuantiTect primer assays (Qiagen) were used in experiments. Primer information is available upon request.

Acknowledgements

We thank K. Kleinhans and A. Lechel for help with microarray; S.F. Katz for FACS-support; P. Mahaddalkar for support with lentivirus; and V. Sakk and J. Nichols for help with embryo stainings; L. Vallier, Eomes-reporter; J. Wells, Sox17-reporter; K. Hochedlinger, lentiviral LVdTomato construct; R. Kemler, T-reporter; D. Melton, Sox17-reporter cells; H. Fehling, Dpp4- RFP/Brachyury-GFP cells; V. Christoffels, TBX3-Reporter mice; and S. Arnold, G. von Wichert, H. Friedle, K. L. Rudolph, and A. Wobus for critically reading the manuscript.

Supplemental References


Research Article

Definitive Endoderm Formation from Plucked Human Hair-Derived Induced Pluripotent Stem Cells and SK Channel Regulation

Anett Illing,1 Marianne Stockmann,1 Narasimha Swamy Telugu,1 Leonhard Linta,2 Ronan Russell,1 Martin Müller,1 Thomas Seufferlein,1 Stefan Liebau,2 and Alexander Kleger1

1 Department of Internal Medicine I, University of Ulm, Albert-Einstein Allee 23, 89081 Ulm, Germany
2 Institute for Anatomy and Cell Biology, University of Ulm, Albert-Einstein Allee 11, 89081 Ulm, Germany

Correspondence should be addressed to Stefan Liebau; stefan.liebau@uni-ulm.de and Alexander Kleger; alexander.kleger@uni-ulm.de

Received 18 January 2013; Accepted 13 March 2013

Academic Editor: Michael Levin

Copyright © 2013 Anett Illing et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pluripotent stem cells present an extraordinary powerful tool to investigate embryonic development in humans. Essentially, they provide a unique platform for dissecting the distinct mechanisms underlying pluripotency and subsequent lineage commitment. Modest information currently exists about the expression and the role of ion channels during human embryogenesis, organ development, and cell fate determination. Of note, small and intermediate conductance, calcium-activated potassium channels have been reported to modify stem cell behaviour and differentiation. These channels are broadly expressed throughout human tissues and are involved in various cellular processes, such as the after-hyperpolarization in excitable cells, and also in differentiation processes. To this end, human induced pluripotent stem cells (hiPSCs) generated from plucked human hair keratinocytes have been exploited in vitro to recapitulate endoderm formation and, concomitantly, used to map the expression of the SK channel (SKCa) subtypes over time. Thus, we report the successful generation of definitive endoderm from hiPSCs of ectodermal origin using a highly reproducible and robust differentiation system. Furthermore, we provide the first evidence that SKCa subtypes are dynamically regulated in the transition from a pluripotent stem cell to a more lineage restricted, endodermal progeny.

1. Introduction

Mammalian development is a tightly regulated process, with considerable biochemical and physiological changes occurring from the time of fertilization to the onset of gastrulation and further differentiation towards fully formed organisms. However, understanding early fate decision events, such as segregation of the three germ layers, is a prerequisite for regenerative medicine [1–5]. The advent of induced pluripotent stem cells and their unique features of unlimited self-renewal and nonrestricted differentiation capacity marked a milestone in the battle to dissect such processes—directly in the context of human development [6–8]. Given the incredible accordance of embryonic development in vivo and its respective model system in vitro, it is not surprising that most of the currently available pluripotent stem cell differentiation protocols make use of physiological, stage-specific signalling clues in order to recapitulate development of all three germ layers: ectoderm, mesoderm, and endoderm. Further differentiation towards more specialized cell types has also been achieved, for example, formation of primitive gut tube endoderm (SOX17/Hnf5b positive [9, 10]), pancreatic progenitor cells (Pdx1/C行政处罚 positive [11, 12]), and hepatic progenitor cells (AFP/HNF4a positive [13]) from definitive endoderm progenitor cells. Nevertheless, the precise mechanisms governing such complex processes are
not completely understood. Another limitation exists in achieving highly homogenous, reproducible cell type-specific yields. As a result, the current use of hiPSCs for disease modelling where the aim is to use in vitro differentiated patient-specific pluripotent stem cells to replace the patients’ damaged cells is massively hindered. In consequence, critically defined, efficient, and robust differentiation protocols are highly anticipated.

Endoderm comprises the innermost of the primary germ layers of an animal embryo and has a primary role to provide the epithelial lining of two major tubes within the body. The first tube, which extends the entire length of the body, is known as the digestive tube and undergoes budding during embryogenesis to form the liver, gallbladder, and pancreas. The second tube, the respiratory tube, forms an outgrowth of the digestive tube and gives rise to the lungs. Notably, two distinct sets of endoderm can be distinguished in the developing embryo: visceral endoderm arising directly from the inner cell mass and definitive endoderm (DE) derived from mesendoderm within the anterior primitive streak in close proximity to the cardiovascular progenitors [1, 14–16]. The visceral endoderm forms the epithelial lining of the yolk sac [17] while the DE is responsible for the internal (mucosal) lining of the embryonic gut and is governed by the expression of key transcription factors such as SOX17 [18], Foxa2, or Hex1 [19].

To date, a large group of proteins has been broadly neglected concerning its role during developmental processes, namely, ion channels. In addition to the modulation of the membrane potential in various tissues and cell populations, ion channels were identified to be involved in a number of biological processes, such as proliferation, cell differentiation, and cell morphogenesis. Since these mechanisms are apparently abundant in the transition of stem or progenitor cell populations to more defined cell types of different origin and potency, a role for ion channels in developmental processes can be hypothesized [20–23]. In particular, the adsorptive tissues derived from the DE are often rich in ion channels and defects in these channels are responsible for some harmful diseases. One prominent example is cystic fibrosis (CF), a common, autosomal recessive disorder due to mutations in a chloride channel known as the CFTR. Located on the plasma membrane of many epithelial cells, this simple mutation gives rise to abnormalities of salt and fluid transport in many endodermal derived tissues including lung, pancreas, and liver [24]. However the contribution of other ion channel families to diseases within the foregut has apparently been poorly studied.

Indeed, in pluripotent stem cells, activation of small and intermediate conductance calcium activated potassium channels (SK channels; SKCas) triggers the MAPK/ERK pathway following RAS/RAF activation finally, giving rise to cytoskeletal rearrangement, cardiogenesis, and cardiac subtype specification [2, 3, 5, 25]. The group consists of four members, namely, SK1 (KCa 2.1, KCNN1), SK2 (KCa 2.2, KCNN2), SK3 (KCa 2.3, KCNN3), and SK4 (KCa 3.1, KCNN4). The functional form of the ion pore is mediated by the combination of the 4 subunits, respectively. Additionally, widely distributed functional splice variants of SKCas have been found throughout the organism in several tissues [26–28]. Functional SKCas are not only constructed as homo- but also as hetero-tetrameric channel proteins, most probably serving a cellular and functional specificity [26, 29]. The pore is opened following subtle elevation of intracellular calcium levels. Calmodulin, attached in a Ca⁡²⁺-dependent manner to the C-terminal of the channel subunits, specifically binds Ca⁡²⁺-ions and mediates a conformational change of the channel protein, leading to the opening of the pore [30, 31]. Calcium is the only known physiological activator of SKCas and channel opening occurs within a few milliseconds [31]. SK1-3 are highly expressed in the nervous system where they modify the membrane potential; that is, they crucially contribute to the after-hyperpolarization and therefore regulate the firing pattern, frequency, and length of action potentials in different neuronal networks [32–35]. On the other hand, SKCas play important roles in multiple other cellular functions, namely in cerebral and peripheral blood vessel smooth muscle, the functional myocardium, or neural progenitor cells [21, 36, 37].

In the current study, we highlight a robust and efficient differentiation protocol to drive plucked human hair-derived hiPSCs towards definitive endoderm. Furthermore, we analyse changes in protein and mRNA expression in the SKCa family of ion channels in the transition from a pluripotent cell state to a definitive endodermal committed cell type.

2. Materials and Methods

2.1. Keratinocyte Culture from Plucked Human Hair. Outgrowth of keratinocytes from plucked human hair was induced as described previously [25, 38]. Keratinocytes were split on 20 μg/mL collagen IV-coated dishes and cultured in EpLife medium with HKGS supplement (both Invitrogen, USA). The use of human material in this study has been approved by the ethical committee of the Ulm University (Nr. 0148/2009) and in compliance with the guidelines of the Federal Government of Germany and the Declaration of Helsinki concerning Ethical Principles for Medical Research Involving Human Subjects.

2.2. Rat Embryonic Fibroblasts (REFs) Culture. REFs were isolated from day E14 Sprague Dawley rat embryos as described previously [38] and cultured in DMEM supplemented with 15% FCS, 2 mM GlutaMAX, 100 μM nonessential amino acids, and 1% Antibiotic-Antimycotic (all Invitrogen). Cells were passaged using 0.125% trypsin digestion when reaching confluence for up to 5 passages. All animal experiments were performed in compliance with the guidelines for the welfare of experimental animals issued by the Federal Government of Germany, the National Institutes of Health, and the Max Planck Society (Nr. O.103).

2.3. Lentivirus Generation. Lentivirus containing a polycistronic expression cassette encoding OCT4, SOX2, KIF4, and cMYC was produced as described previously [3, 25, 38].
2.4. Reprogramming Keratinocytes. Keratinocytes at 75% confluence were infected with 5 × 10⁵ proviral genome copies in EpilLife medium supplemented with 8 μg/mL polybrene on two subsequent days. On the third day, keratinocytes were transferred onto irradiated REF feeder cells (2.5 × 10⁵ cells per well irradiated with 30 Gy). Cells were cultured in hiPSCs medium in a 5% O₂ incubator and medium was changed daily. After 3–5 days small colonies appeared, showing a typical hiPSCs like morphology. Around 14 days later, hiPSC colonies had the appropriate size for mechanically passaging and were transferred onto irradiated MEFs or onto Matrigel-coated (BD, USA) dishes for further passaging.

2.5. hiPSC Culture. hiPSCs were initially cultured on feeder cells in Knockout DMEM (Invitrogen) supplemented with 20% Knockout Serum Replacement (Invitrogen), 2 mM GlutaMAX (Invitrogen), 100 μM nonessential amino acids (Invitrogen), 1% Antibiotic-Antimycotic (Invitrogen), 100 μM β-mercaptoethanol (Millipore, USA), 50 μg/mL ascorbic acid, and 10 ng/mL FGF2 (both PeproTech, USA) in a 5% CO₂ incubator.

For later passages, hiPSCs were mechanically picked and transferred onto Matrigel-coated dishes (BD) and kept in FDTA medium that contains DMEM/F12-GlutaMax (Invitrogen), 1 × ITS (Insulin-Transferrin-Selenium, Invitrogen), 0.1% HSA (Biological Industries, Israel), 1 × Lipid mix (Invitrogen), 1 × Penicillin-Streptomycin (Millipore), 10 ng/mL FGF2 (PeproTech), 0.5 ng/mL TGF-β (PeproTech), 50 nM Dorsomorphin (Sigma, Germany), and 5 ng/mL Activin A (PeproTech) [39]. hiPSCs were cultured in a 5% O₂ incubator and medium was changed daily. For splitting cells, hiPSC colonies were incubated with Dispase (StemCell Technologies, France) for 5–7 min at 37°C and subsequently detached using a cell scraper. After brief centrifugation, cells were resuspended in FDTA medium and transferred onto Matrigel-coated dishes. Excision of the viral cassette was achieved by incubation with recombiant TAT-Cre protein (1.5 μM for 5 h).

2.6. Monolayer-Based hiPSC Differentiation towards Definitive Endoderm. For DE differentiation, hiPSCs were plated onto Matrigel-coated dishes and cultured in FDTA medium supplemented with 10 μM Rock-inhibitor Y-276342 (Ascent, UK) for 24 hours. When the cells reached about 75% confluence, medium was changed to RPMI 1640 medium (Invitrogen) containing 2% FBS (Lonza, CH) with 500 nM IDE1 (R&D systems, USA), 3 μM CHIR99021 (Axonemedchem), 5 μM LY294002 (Sigma, Germany), and 10 ng/mL BMP4 (PeproTech) for 24 hours. Then, medium was changed to RPMI 1640 medium containing 2% FBS and supplemented with 500 nM IDE1 and 5 μL LY294002 for two days. From day 3 on, cells were cultured in RPMI 1640 supplemented with 500 nm IDE1, 5 μM LY294002, and 50 ng/mL FGF2. The respective figure contains an experimental outline illustrating detailed culture conditions and treatment regimens [40–45] (Figure 2a).

2.7. Immunocytochemistry. Immunofluorescence has been previously described [5]. Nuclei were stained with DAPI. Primary antibodies were used as follows: SKI/2 (both 1:100), SK3 (1:100, Alomone Labs, Israel), SK4 (1:100, Cell Applications, USA), SOXI7 (1:500, R&D systems), and FOX2A (1:100, Santa Cruz, USA). hiPSC lines were characterized using the StemLife Pluripotency Antibody Kit (Cell Signaling, USA). Fluorescence labelled secondary antibodies were Alexa Fluor 488 and Alexa Fluor 568 (both Invitrogen). Images were captured using an upright fluorescence microscope (Axioskop 2, Zeiss, Oberkochen, Germany) equipped with a Zeiss CCD camera and analysed using Axiovision software (Zeiss) [46].

2.8. Quantitative One-Step Real Time. RT-PCR (qPCR) Analysis was performed as previously described. Briefly, one-step real-time qPCR was carried out with the LightCycler System (Roche, Mannheim, Germany) using the QuantiTect SYBR Green RT-PCR kit (Qiagen, Hilden, Germany). Relative transcript expression was expressed as the ratio of target gene concentration to the housekeeping gene hydroxymethylbilane synthase (HMBS) [47, 48].

2.9. FACS Analysis. For flow cytometry cells were harvested with TrypLE (Invitrogen) for 7 min at 37°C to obtain single cells suspension. Next, cells were washed twice with PBS, blocked with 5% HSA-solution (in PBS) to avoid unspecific binding of the antibodies to the Fc-receptor. Cells were washed again with PBS and incubated for 40 min at 4°C with CXCR4-PE (Invitrogen), subsequently c-Kit-APC (Invitrogen) was added for additional 10 min at 4°C in FACS buffer (2% FCS in PBS), according to the manufacturer’s instructions. Cells were washed with FACS buffer, 50 ng/mL DAPI was added, to exclude dead cells from analysis, and the samples were directly analysed on an LSRII flow cytometer (BD).

For intracellular SOX17 staining cells were washed twice with PBS, blocked with 5% HSA-solution (in PBS) to avoid unspecific binding of the antibodies to the Fc-receptor. Cells were washed again with PBS and the pellet was resuspended in 4% PFA and incubated for 15 min at 37°C for fixation. Subsequently the cell pellet was resuspended in 0.5% Saponin in FACS buffer (saponin buffer) and incubated for 30 min on ice. Cells were pelleted and stained with SOX17 (1:100, R&D systems) at 4°C for one hour. Cells were washed with saponin buffer and afterwards incubated for 30 min at 4°C with anti-goat Alexa Fluor 647. Finally cells were washed with FACS buffer and directly analysed on an LSRII flow cytometer (BD).

2.10. Statistical Analysis. If not stated otherwise, error bars indicate standard deviations. Calculations were done with GraphPad Prism 5 (GraphPad Software, Inc., San Diego, http://www.graphpad.com/).

3. Results
3.1. Reprogramming Human Hair-Derived Keratinocytes to hiPSCs. For the depicted studies we utilized keratinocyte
Figure 1: Continued.
cultures from plucked human hair of healthy individuals (Figure 1(a)). With the use of a lentiviral, multicistronic four-factor reprogramming system, keratinocytes were successfully reprogrammed to human induced pluripotent stem cells displaying embryonic stem cell like morphology (Figure 1(a)) as well as hallmarks of pluripotency tested via immunohistochemistry and qRT-PCR for the expression of embryonic stem cell markers. Several lines of more than 5 individuals (data not shown or reported in [3, 25, 38]) were tested for their proliferation and differentiation capacity and subsequently two lines were selected, named “hiPSC_1 and hiPSC_2.” Both lines were additionally tested for the protein expression of OCT4, SOX2, NANOG, SSEA4, TRA1-60, and TRA1-81 (Figure 1(b)) and mRNA levels of three pluripotency markers (OCT4, SOX2, NANOG). At the pluripotent stage definitive endoderm makers (FOXA2, SOX17) and markers for pancreatic progenitors (PTFA1, PDX1) were negative (Figure 1(d)). Additionally, all lines were capable of differentiating into cells of all 3 germ-layers, as shown by β-3-tubulin (neurons—ectoderm), α-actinin (muscle cells—mesoderm), and α-fetoprotein (liver cells—endoderm) (Figure 1(c)). One line was further treated with recombinant Cre protein to excise the reprogramming STEMCCA cassette being flanked with loxp sites. To test for successful excision, PCR amplification of the STEMCCA cassette was performed in cre- and nontreated iPS cell clones from this respective line showing the band only in controls (Figure 1(e)). Taken together, our established hiPSC lines display an embryonic stem cell like phenotype, proven by morphology and expression of pluripotency markers as well as absent mRNA for endodermal markers.

3.2. Human Induced Pluripotent Stem Cells Can Be Differen-
tiated to Cells Representing the Definitive Endoderm. To test the differentiation potential of our established hiPSC lines into definitive endoderm (DE), previously published protocols were combined in terms of a small molecule-driven
**Figure 2: Continued.**
Figure 2: Differentiation of hiPSCs into DE (definitive endoderm) cells. (a) Scheme of monolayer differentiation of hiPSCs into DE cells. (b) Immunocytochemistry shows that hiPSCs-derived DE cells express the early endodermal markers SOX17 (red) or FOXA2 (green). Scale bars as indicated. (c) FACS analysis reveals that after 3 days of differentiation approximately 45% of the cells have become SOX17 positive. After 6 days of differentiation about 80% of the cells express SOX17. This data was consistent within two different hiPSC lines. (d) Number of CXCR4/cKIT double-positive cells after 3 days and 5 days of differentiation, respectively. After 5 to 6 days of differentiation both hiPSCs lines express the highest amount of CXCR4/cKIT double-positive cells in the region of 75–80%. (f) Virus-containing hiPSCs (hiPSC_Cre) did not show differences in the number of CXCR4/cKIT double-positive cells compared to virus-free iPSCs (hiPSC_Cre). (g) Continuous loss of OCT4 mRNA levels during DE differentiation. In contrast, transcript levels of early endodermal genes such as SOX17 and FOXA2 steadily increased and reached highest levels after 5 days of differentiation. Expression levels are shown relative to the housekeeping gene HMBS (n = 4, two different hiPSC lines).

approach [43–45]. Small molecule-based assays are less biased by batch-to-batch variations and are usually more cost effective. Upon extensive testing of different combinations, our protocol led to the following replacements of established growth factors being known to drive definitive endoderm differentiation: CHIR90021 replaced Wnt3a [40], IDE1 replaced Activin A [41], and LY294002 inhibited the AKT signalling pathway to abolish pluripotency [42]. Figure 2(a) represents a detailed scheme of the differentiation conditions used for the formation of DE from day 0 (undifferentiated pluripotent hiPSCs) to day 6 (definitive endodermal cells). In vitro differentiated hiPSCs became positive for endodermal markers, confirmed by positive immunostaining of cells on day 5 for FOXA2 and SOX17 (Figure 2(b)). To analyse and characterize the SOX17 expression more objectively, we quantified SOX17 expression via intracellular FACS analysis in a time course from day 3 to 6 of the protocol. Figure 2(c) represents representative FACS plots from both lines, representing SOX17 positive cells on day 3 and day 5. We did not observe differences in the differentiation capacity of virus-free hiPSCs after excision of the reprogramming cassette in comparison to virus-containing cells, making further analyses of silencing of exogenous factors unnecessary (Figure 2(d)). In summary, SOX17 expression is increasing from approximately 45% at day 3 to nearly 80% of SOX17 positive cells on day 6. Recent publications depict CXCR4 and c-KIT positive cells as definitive endoderm progenitors, that give rise to self-renewing endodermal progenitor cells (EPCs) [49]. To confirm our protocol, we did time course analysis by flow cytometry for CXCR4 and c-KIT positive cells during differentiation. Figure 2(e) shows representative FACS plots of CXCR4 and c-KIT positive cells of the two hiPSC lines on day 3 and 5. Two independent experiments for each line were summarized and shown from day 2 to 7 of endodermal differentiation. From day 2 on, the double positive population (CXCR4 and c-KIT) is steadily increasing in both lines. The maximum is reached with almost 90% double positive cells for hiPSC_1 and almost 80% for hiPSC_2 (Figure 2(d)). Again there was no relevant difference upon excision of the reprogramming cassette (Figure 2(f)).

To further confirm the definitive endodermal identity of the differentiated lines, we measured mRNA levels using qRT-PCR analysis for OCT4, SOX17, and FOXA2. From day 1 to day 5 mRNA levels for the pluripotency marker OCT4 decrease continuously (Figure 2(g), summarized for hiPSC_1 and hiPSC_2). SOX17 and FOXA2 levels were tested in the two established hiPSC lines during differentiation and displayed increasing mRNA levels from day 1 to day 5 (Figure 2(g)). This data clearly indicates that the investigated hiPSC lines can be differentiated into DE, loosing markers of pluripotency and up regulating the expression of endodermal markers during endoderm formation.

3.3. Expression of Calcium-Activated Potassium Channels (SKCas) during DE Differentiation. Next, we had a closer look on the expression of the different SKCas subtypes during DE differentiation. hiPSCs were differentiated into DE cells and expression of SKCa was investigated after 5 days of differentiation. On day 5 SOX17 is strongly expressed indicating the differentiation into DE cells (Figure 3(a)). To analyse the expression of the SKCas, DE cells were stained for the different SKCa subtypes. Immunofluorescence analyses show a quite strong and stable expression of SK1, SK2, and SK3 whereas SK4 seems to be expressed at a lower level (Figure 3(b)). SK1, 2, and 4 are localized in the cytoplasm and the cell membrane. However, SK3 is not only localized at the cell membrane but also as PUNCTUA
Figure 3: Expression of Calcium-activated Potassium channels during formation of definitive endoderm. (a) Expression of SOX17 (green) after 5 days of DE differentiation. (b) Immunofluorescence analysis of SKCa proteins in DE cells. Indicated SKCa subtype (red). Scale bars as indicated. (c) Higher magnifications of indicated SKCa subtype (red). Scale bars as indicated. (d) Transcript levels of SK1 and SK2 remained relatively low during the DE differentiation. In contrast, mRNA levels of SK3 increased after 4 days of differentiation. SK4 mRNA levels slightly increased during the first days of differentiation and peaked on day 3 followed by a sharp decrease until day 5. Expression levels are shown relative to the housekeeping gene HMBS (n = 4, two different hiPSCs lines).
in the nuclei (Figure 3(c)). This is a finding that needs to be analysed in further studies. Double immunofluorescence staining for SOX17 and respective channel proteins are shown in Supplementary Figure 1 available online at http://dx.doi.org/10.1159/2013/360573.

mRNA expression analysis via quantitative RT-PCR (qRT-PCR) shows a relative constant expression of SK1 and SK2 during DE differentiation (Figure 3(d)). In contrast, transcript levels of SK3 obviously increased after 4 days of differentiation (Figure 3(d)). SK4 mRNA levels marginally increased during the first days of differentiation and peaked on day 3, followed by a sharp decline up to day 5 (Figure 3(d)). To note, all four SKCa subtypes are expressed during DE differentiation. SK1 and SK2 are constantly expressed whereas SK3 seems to be up regulated during ongoing DE differentiation. The transcript levels of the different SKCa subtypes on day 5 reflect our observations of the immunofluorescence analysis. In sum, all 4 SK subtypes are differentially expressed during DE differentiation of human induced pluripotent stem cells with a yet undescribed localization of SK3 in the nucleus.

4. Discussion

In the current study, we provide proof of the concept that plucked human hair-derived iPSCs are highly potent in their capacity to commit not only towards mesoderm [3] and neuroectoderm [25] but also towards the endodermal germ layer, particularly definitive endoderm. To this end, a newly adopted protocol based on previously published studies was applied and resulting cells were extensively characterized by gene expression analysis, immunofluorescence microscopy, and FACS-staining for intracellular and surface markers defining the definitive endoderm signature.

As induced pluripotent cells are currently considered to resemble human embryonic stem cells, a state-of-the-art assay for hiPSC generation is required. Such an assay requires the following prerequisites: (i) noninvasive harvest of the cell type of origin, (ii) broad applicability in terms of guided differentiation to all three germ layers, (iii) useful for large-scale hiPSC biobanking, (iv) highly efficient, and (v) fast reprogramming to the hiPSC stage. Keratinocytes from the outer root sheath of plucked human hair represent such a cell source and thus points towards the generation of patient-specific human induced pluripotent stem cells as a new paradigm for modelling human disease and for individualizing drug testing. Previously, we have further optimized this method in terms of efficiency and speed by using rat embryonic fibroblasts as a feeder layer for keratinocyte reprogramming [38]. The arising hiPSCs fulfilled all the prerequisites of pluripotency including teratoma formation and spontaneous three-germ layer differentiation.

In further studies, we have applied plucked hair-derived hiPSCs to guide differentiation towards motoneurons [25] and cardiac pacemaker cells [3], both representing highly specified cell types from either ectodermal or mesodermal origin. However, their differentiation capacity to give rise to definitive and primitive gut tube endoderm remained elusive. While forming, definitive endoderm is incorporated by morphogenetic movements into a primitive gut tube stage. This in turn is patterned into foregut, midgut, and hindgut to form the functional epithelial compartment of multiple internal organs: liver, intestines, lungs, and the pancreas [50]. Nowadays, virtually every cell population arising from the primitive gut tube has been generated using guided differentiation of pluripotent cells towards liver, intestines, lungs, and the pancreas [51–53]. Thus, the induction of DE cells marks a prerequisite for the entire process of pluripotent stem cell differentiation into, for example, pancreatic or hepatic progenitor cells [54, 55]. Several protocols have been developed and modified to increase the efficiency of DE commitment. All these protocols are strongly dependent on high doses of TGFβ signalling mediated by Activin A as the major driving force of the process. However, large-scale differentiation experiments should be cost effective, thus, making a small molecule-based assay more desirable. To this end, we combined several previously described strategies. First, we replaced Activin A by IDE1, a compound having shown to display similar but also superior characteristics compared to Nodal or Activin A [41]. Similarly, we substituted Wnt3a by the small molecule CHIR99021 that inhibits GSK 3 kinase to mimic Wnt signalling [40]. The third small molecule LY294002 inhibited the AKT signalling pathway, by repressing PI3K, to promote the exit from pluripotency [42]. In consequence, a robust and reproducible assay was developed which shows to be effective in several human plucked hair-derived iPSCs. As the formation of definitive endoderm is a prerequisite to obtain, for example, relevant numbers of pancreatic β-cells, our data in combination with the presented reprogramming strategy are highly relevant for human disease modelling approaches.

However, several studies have suggested that β-cells generated from human pluripotent stem cells lack adult, and at the most reach, fetal maturity as particularly expressed by their polyhormonality. This observation reinforces the notion that establishing culture conditions that promote appropriate maturation represents a significant obstacle for the generation of functional β-cells in vitro [56]. A recent landmark paper identified self-renewing definitive endodermal progenitor cells as a potential cell source to bypass this limitation. β-cells generated from these cells showed features of adult maturity as even shown by functional assays [57]. Given the fact that all published protocols so far lack this feature, the quality of the definitive endodermal intermediate seems to have an impact on the final maturity. The generation of definitive endodermal progenitor cells was characterized by high positivity for c-KIT and CXCR4 [57]. Thus, we included in our current DE analysis an FACS-based tool and indeed succeeded in obtaining a pattern likely to allow the isolation of this distinct cell type. The similar differentiation capacity of all our analysed plucked human hair-derived iPSCs is relevant to the field of disease modelling, using patient-specific material. Plucked hair keratinocytes are more or less the only cell type, which matches the above criteria. Nevertheless, a potential ectodermally biased epigenetic memory could limit their utility [58]. Our finding abolishes such a bias at least based on the number of different cell lines and the reproducible endodermal commitment pattern.
The development of in vitro models underlying embryonic development is a prerequisite to build new knowledge and to develop new strategies targeting various genetic diseases. The development and investigation of endoderm-derived cells are such as pancreatic cells, are of high importance for the field of developmental biology and clinical implications. Induced pluripotent stem cells (iPSCs) with their unique features of unlimited self-renewal and non-restricted differentiation capacity are a highly promising tool for regenerative medicine as well as for studies on developmental biology. iPSCs have been generated from a variety of different cell types originating from all three germ layers [38, 58, 59]. Finally, this setup has been used to determine the expression pattern of a certain ion channel family which has been previously shown to be differentially regulated in embryonic stem cells and involved in differentiation processes, namely, small and intermediate conductance calcium-activated potassium channels [2, 3, 5, 60]. Thus, our study gives novel insights into guided pluripotent stem cell differentiation towards definitive endoderm and a potentially involved protein family.

SKCas either exhibit small (SK1, KCa2.1, Kcnnl; SK2, KCa2.2, Kcnn2; and SK3, KCa2.3, Kcnn3) or intermediate (SK4, K, KCa3.1, and Kcn4) unitary conductance for K⁺ ions. Important roles in multiple cellular functions, for example, cell cycle regulation in cancer cells [20, 61], smooth muscle relaxation [23, 62], mesenchymal stem cell proliferation [22], and cytoskeleton reorganization in neural progenitors [21] have been reported. SKCas are widely expressed throughout all different tissues. While SK1 is exclusively expressed in the central nervous system, SK2 is more widely expressed in different organs arising from different germ layers such as brain, liver, or heart. SK3 is the most widespread expressed isoform with a predominant expression pattern in the central nervous system but also in smooth muscle rich tissues. SK4 can be detected in inflammatory cell-rich, surface-rich, and secretory tissues such as the pancreas [63]. In the pancreas, for example, SK4 regulates glucose homeostasis and enzyme secretion of acinar cells [64, 65]. Moreover, SKCas are overexpressed in a variety of cancers, including pancreatic cancer [20] and, for example, SK3 was shown to be involved in cancer cell migration [66]. Nevertheless, the role of SKCas in developmental processes remains enigmatic though it is well accepted that cell differentiation and maturation affect the expression patterns of ion channels. Our group has shed for the first time light on their role in differentiating pluripotent stem cells derived from mouse and men [2, 3, 5, 25]. A potential role of SKCas was already suggested by their differentially regulated expression pattern. In fact, it temporally coincides with the commitment of the cardiovascular progenitor showing an expression peak of the respective isoform around day 5 [5]. In consequence, we aimed to address the expression pattern of SKCas in the developing endoderm using plucked human hair-derived iPSCs as a bona fide modelling system. Interestingly, the differential regulation of most SKCa isoforms was relatively modest. Albeit SK2 and SK4 show a slight expression peak around day 2/3, the only regulated isoform seems to be SK3 showing a continuously increasing expression with ongoing DE formation. Interestingly, reports showing SK3 expression in DE-derived organs are restricted to a handful of studies showing SK3 expression in epithelial cancer cells and a liver-specific splice variant [67, 68]. Further studies including gain and loss of function approaches within the same assay have to clarify the respective functions of SKCa isoforms with DE formation and later maturation processes towards liver and pancreas.

To summarize, we present an efficient, novel, and robust DE formation assay being suitable for ectoderm-derived plucked human hair iPSCs. Given the prerequisites for reprogramming fulfilled by plucked human hair, a robust DE assay for this particular iPSCs type is highly relevant for disease modelling approaches. Subsequently, we have identified dynamic expression of the SKCa family of proteins during DE formation.

Conflict of Interests

The authors declare no conflict of interests.

Authors’ Contribution

Anett Illing and Marianne Stockmann contributed equally.

Acknowledgments

The authors thank Stefanie Fischer, Ralf Köhntop, and Sabine Seltenheim for the excellent technical assistance enabling research to be conducted in our laboratories. The authors thank all members of the Kleger and Liebau lab for discussing and critically reading the paper. The research that enabled this paper was funded by a Fellowship by the Medical Faculty of Ulm University (Bausteinprogramm, L. SBR.0011) and by the Deutsche Forschungsgemeinschaft (KL 2544/1-1, SL BO1718/4-1), the German Foundation for Heart Research (F/34/II; to AK and SL), Boehringer-Ingelheim BIU (N5 to SL; C1 to AK), the virtual Helmholtz-Institute (VHI-V-510 to SL), and the Else-Kröner-Fresenius-Stiftung (2011_A200; to AK and SL). AK is indebted to the Baden-Württemberg Stiftung for the financial support of this research project by the Eliteprogramme for Postdocs. The authors apologize to authors whose work has not been covered or directly cited due to space limitations.

References


[35] O. Waroux, L. Massotte, L. Alleva et al., "SK channels control the firing pattern of midbrain dopaminergic neurons in vivo,"


M. Potier, V. Joulin, S. Roger et al., "Identification of SK3 channel as a new mediator of breast cancer cell migration,"


Supplementary Figure 1 Immunofluorescence analysis of Calcium-activated Potassium Channels during definitive endoderm differentiation. (A, B) Expression of SOX17 (green) and SK3 or SK4 (red) after 5 days of DE differentiation.
Loss of ATM accelerates pancreatic cancer formation and epithelial-mesenchymal transition

Ronan Russell¹, Lukas Perkhofer¹,* , Stefan Liebau²,* , Qiong Lin³, André Lechel¹, Fenja M. Feld⁴, Elisabeth Hessmann⁵, Jochen Gaedcke⁶, Melanie Güthle¹, Martin Zenke³, Daniel Hartmann⁷, Guido von Figura⁸, Stephanie E. Weissinger⁴, Karl-Lenhard Rudolph⁹, Peter Möller⁴, Jochen K. Lennerz⁴, Thomas Seufferlein¹, Martin Wagner¹,† & Alexander Kleger¹,†

Pancreatic ductal adenocarcinoma (PDAC) is associated with accumulation of particular oncogenic mutations and recent genetic sequencing studies have identified ataxia telangiectasia-mutated (ATM) mutations in PDAC cohorts. Here we report that conditional deletion of ATM in a mouse model of PDAC induces a greater number of proliferative precursor lesions coupled with a pronounced fibrotic reaction. ATM-targeted mice display altered TGFβ-superfamily signalling and enhanced epithelial-to-mesenchymal transition (EMT) coupled with shortened survival. Notably, our mouse model recapitulates many features of more aggressive human PDAC subtypes. Particularly, we report that low expression of ATM predicts EMT, a gene signature specific for Bmp4 signalling and poor prognosis in human PDAC. Our data suggest an intimate link between ATM expression and pancreatic cancer progression in mice and men.
Despite intensive basic and clinical research, deaths due to pancreatic ductal adenocarcinoma (PDAC) rank fourth among cancer-related events in the western world, with an overall 5-year survival rate around 4% (ref. 1). Genetically engineered animal models effectively recapitulate both the morphological and molecular features of PDAC and have helped to identify key factors within the genetic landscape directing PDAC formation such as oncogenic K-ras, Trp53 and Ink4A/Arf. This has led to a better understanding of the molecular mechanisms that drive PDAC. Genome-wide exome-sequencing studies recently identified several novel mutations associated with PDAC2–5. However, the precise contribution of these newly identified factors within tumour biology remains elusive.

Ataxia telangiectasia mutated (ATM) is a serine/threonine kinase and was initially characterized for its role in the DNA damage response (DDR)6–8. Interestingly, accumulating evidence suggests that ATM has a broader capacity to integrate and direct various signalling cues to maintain cellular homeostasis than previously appreciated9,10. These include regulation of chromatin remodelling, oxidative stress, and cellular metabolism in diverse tissues. Patients with the recessive disease ataxia telangiectasia (AT) and ATM-deficient mice exhibit immunodeficiency, genomic instability, and an increased risk for lymphoid malignancies. Moreover, ATM deficiency has an impact on self-renewal of hematopoietic stem cells, accelerates ageing in telomere dysfunctional mice, and drives angiogenesis9. Thus, it is becoming increasingly clear that ATM serves multiple functions in a variety of cellular compartments. To date, the role of ATM in pancreatic cancer initiation/progression is largely unclear. Inactivating variants of the ATM gene are carried by ∼1% of the general population11 and more recently, familial pancreatic cancer patients have been shown to harbour a nonsense ATM germ line mutation leading to somatic loss of the variant allele12,13. Recent data based on large-scale sequencing studies reported up to 18% of ATM mutations in certain human PDAC cohorts, which can be even detected in the germ line of certain individuals3,5,14–16. In line, a recent report underscored this finding on the protein level17. This observation positions ATM next to established factors such as K-ras or Trp53 among the 16 most commonly mutated genes in PDAC5,15,16.

Herein we examine the role of ATM in PDAC tumour biology in both mice and men and provide evidence that loss of ATM (1) enhances acinar-to-ductal reprogramming (ADR) via altered TGFβ-superfamily signalling, (2) is associated with epithelial-to-mesenchymal transition (EMT) and a gain in tumour initiating properties and (3) acts as an independent prognostic marker as ATM depletion correlates significantly with survival.

**Results**

**Loss of ATM promotes ADR and neoplastic lesion formation.**

First, we reanalysed currently available information on ATM mutations from two comprehensive PDAC data sets from the International Cancer Genome Consortium. These independent cohorts from Canada and Australia show mutation frequencies of 9–18% averaging to about 12% within the ATM gene. This includes non-synonymous substitutions and insertion-deletions (indels) and in line with recent reports, underscores the clinical and biological relevance of ATM mutations in a significant subset of human pancreatic cancers (Table 1)18–22.

To examine the role of ATM deficiency in pancreatic carcinogenesis, we crossed mice harbouring a floxed Atm (A′) allele to an existing PDAC model, p48CreER+/− (C′), KrasG12D+/− (K′) mice (Fig. 1a)18. Animals containing all three alleles are referred to as AKC mice throughout the text and characterization of target animals is described in detail in Supplementary Fig. 1a–d. Of note, animals with loss of ATM alone showed normal pancreatic development at all time points examined (Supplementary Fig. 1e). Initial analysis on mice at 5 weeks of age revealed a small number of foci with altered acinar architecture within the pancreas of KC mice. In contrast, age-matched AKC animals already showed more parenchymal foci with disruption of acinar tissue (Fig. 1b).

The early stages of human PDAC are characterized by the onset of defined ductal precursor lesions, the so-called acinar-to-ductal metaplasias (ADMs), acinar-to-ductal reprogramming (ADRs) and pancreatic intraepithelial neoplasias (PanINs)18,19. At 10 weeks of age, we observed more progressive loss of Amylase+ acinar tissue accompanied by an increase in Cytokeratin 19 (CK 19)+ or alcian blue+ ductal precursor lesions in AKC mice compared with KC controls (Fig. 1c,d). Carboxypeptidase (CPA) and SOX9 immunostaining confirmed this finding (Supplementary Fig. 2a). Specifically, we observed a predominant increase of ADM lesions (Fig. 1d,h and Supplementary Fig. 2a) and low-grade PanINs but also increased numbers of high-grade PanINs in AKC mice compared with KC mice (Fig. 1e–g,i). Of note, the most striking changes in pancreatic architecture occurred in the head and upper-body region of the pancreas, while the lower body and tail was generally undisturbed. As previously described for ATM, the phenotype of heterozygous- and homozygous-depleted ATM animals was similar, indicating that haploinsufficiency is sufficient to cause a pancreatic phenotype20–23.

ATM is a well-known cell cycle checkpoint24. Therefore, we checked whether there is any change in the expression of markers and known regulators of the cell cycle. Immunohistochemistry (IHC) for Ki-67, CyclinD1 and CyclinE revealed a higher number of proliferating cells within the ductal precursor lesions of ATM-targeted mice (Fig. 1j,m and Supplementary Fig. 2b,c) compared with controls. Furthermore, Masson–Goldner staining revealed a significant amount of extracellular matrix production particularly in areas of ADM lesions. Intriguingly, ADM lesions of AKC mice showed more α-SMA staining within the areas of neoplastic lesions (Fig. 1k,l,n), indicative of activated pancreatic stellate cells as the major source of stromal infiltration.

**Table 1** | **ATM genetic alterations in human PDAC.**

<table>
<thead>
<tr>
<th>ICGC data set</th>
<th>Number of donors harbouring ATM mutations*</th>
<th>Number of donors in data set</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>PACA-AU</td>
<td>37</td>
<td>392</td>
<td>9.4%</td>
</tr>
<tr>
<td>PACA-CA</td>
<td>36</td>
<td>199</td>
<td>18.1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>591</strong></td>
<td><strong>12%</strong></td>
</tr>
</tbody>
</table>

ATM—ataxia telangiectasia-mutated; AU—Australia; CA—Canada; ICGC—International Cancer Genome Consortium (last accessed: 3/24/2015); PACA—pancreatic cancer; PDAC—pancreatic ductal adenocarcinoma; Mutation frequency of ATM mutations in PDAC. Data extracted from the ICGC database. *Non-synonymous substitutions or indels.

Loss of ATM enhances EMT and stemness. ATM is a central mediator of the DNA-damage response (DDR), which acts as a barrier against tumour progression24. Interestingly, downstream targets of ATM (such as p21 and γH2AX) as well as resulting senescence were not appreciably altered in AKC mice compared with KC mice, while levels of p53 were increased within the precursor lesions of AKC mice (Supplementary Fig. 2d–f). Levels of apoptosis and numbers of inflammatory cells also did not differ (Supplementary Fig. 2g,h).

To gain further insights into the molecular changes of ATM depletion, we performed comparative genome-wide transcriptional profiling followed by gene set enrichment analysis. Using...
Figure 1 | ATM loss of function promotes neoplastic changes in the pancreas in the context of oncogenic K-ras. (a) Illustration of strategy to generate p48Cre;KrasG12D/+;Atm−/− mice (p48Cre/+ = "C"; KrasG12D/+ = "K"; Atm−/− = "A"). (b,c) Representative haematoxylin and eosin (H&E)-stained sections of pancreas at the indicated time points (b) scale bar, 200 μm, (c) scale bar, 500 μm. (d) Immunofluorescence staining of pancreas from the respective genotypes at 10 weeks old shows expression of CK19 (red), amylase (green) and nuclei (Dapi-blue). Scale bar, 20 μm. (e,f) Immunohistochemistry shows expression of (e) CK19 (scale bar, 200 μm) and (f) alcian blue (scale bar, 500 μm) in precursor lesions. Scale bar, 20 μm. (g-i) Quantification of (g) CK19-positive cells, (h) ADM events per visual field, (i) PanIN grading and numbers are shown according to the genotype at 10 weeks. Colour code: Black = Atm+/+; Blue = Atm+/− and Red = Atm−/−. (j-l) Immunohistochemical staining reveals (j) Ki67 (scale bar, 500 μm), (k) fibrosis (Masson-Goldner) (scale bar, 500 μm) and (l) α-SMA (scale bar, 20 μm) at sites of pre-malignant lesions in pancreata from the indicated genotypes and respective quantifications. (m,n). Representative images from at least three mice per genotype are shown. *P<0.05, **P<0.01, ***P<0.0001. One-way analysis of variance (ANOVA). Error bars are the means ± s.e.m.
age-matched (~10 weeks) KC and AKC mice a set of 2,472 differentially regulated genes was identified between the two groups, many of which have previously been implicated in PDAC (Fig. 2a). In addition, we identified a strong deregulation among collagen and matrix-metalloproteinase (MMP) family members (Supplementary Fig. 3a,b) in line with our observation of a pronounced desmoplastic reaction. Sox9 acts as the driving force of acinar-to-ductal metaplasia (ADM) and acinar-to-ductal reprogramming (ADR) in the pancreas15,16. Consistent with this, we found a pronounced upregulation of Sox9 in ATM-depleted mice, in morphologically intact acinar tissue and also in areas of ADM/ADR pointing towards enhanced ductal programming of acinar structures (Fig. 2b,c and Supplementary Fig. 2a).

Certain types of cancers undergo epithelial-to-mesenchymal transition (EMT), a process highly abundant during embryonic development. Interestingly, EMT is particularly prevalent in areas of ADM15 and many factors associated with desmoplasia induce EMT-associated transcription factors (EMT-TFs) such as Sox9 or SLUG26,27. To assess on-going EMT in our model, gene set enrichment analysis (GSEA) of differentially regulated genes between AKC and KC mice was plotted against published gene sets that are implicated in the EMT process28,29. We found that AKC mice present a robust gene signature associated with EMT, as identified using two independent ‘GSEA-core’ sets (Fig. 2d, Supplementary Fig. 3e,f). Of note, 10 random gene sets each containing 100 genes randomly selected from the mouse genome were subjected to GSEA and all the P values are above 0.05, suggesting no similarity or bias to any group of samples (Supplementary Fig. 3d,e). Next, we corroborated the GSEA data via qPCR analysis for Fibronectin, Twist1, Fsp1 and Vimentin (Fig. 2e). Further evidence for increased EMT in precursor lesions of AKC mice is given by IHC for Vimentin and Fibronectin and immunofluorescence staining for ZEB1 (Fig. 2f–h). Taken together, this implies that a significant proportion of ductal epithelial cells in the pancreas of AKC animals have acquired mesenchymal features compared to age-matched KC mice.

To further substantiate these findings and to demonstrate that the observed increase in mesenchymal features is not solely due to increased desmoplasia within the AKC mice, we crossed the AKC mice with Rosa26-tdreFP mice30. This allowed us to trace the cells in which recombination took place and essentially provide a system whereby the epithelial cells of the pancreas could be identified clearly from the surrounding stroma. IHC analysis for the RFP in combination with Zeb1 provides qualitative evidence that indeed the neoplastic lesions, which originate from the epithelium of the pancreas, undergo EMT (Fig. 2i, Supplementary Fig. 4a,b).

To provide a quantitative analysis for EMT between KC and AKC mice, we adopted a strategy that allowed us to isolate a purified population of ductal and neoplastic ductal cells using a DBA-lectin sorting as previously reported31. We first validated this approach as shown in Supplementary Fig. 4d, whereby ductal components, which express Ck19, are enriched in the DBA+ population. Subsequently, by depleting the contaminating stroma and other cell types and enriching for the ductal components, we could confirm that indeed neoplastic duct-like cells of the AKC mice express more markers of EMT (Supplementary Fig. 4c,e). Finally, using micro-dissection techniques to specifically isolate the different neoplastic compartments within the pancreata at 18 weeks of age, we validated our previous observations that AKC mice exhibit more EMT than KC mice. As shown in Fig. 2j, both ADM and PanIN lesions express higher mRNA levels of the EMT-associated genes Sox9, Slug and N-cadherin.

Recently, it has been demonstrated that the acquisition of EMT is associated with cancer stem cells (CSCs), which contribute to the progression of human cancers32. Similar to embryonic stem cells, CSCs exhibit self-renewal and extended differentiation capacity but are also thought to represent the major source of migratory cells with invasive potential. Interestingly, the presence of a CSC gene expression signature in colorectal cancer strongly predicts poor patient’s survival33,34. In line with on-going EMT, GSEA revealed an over-representation of stem cell-associated genes in AKC mice (Fig. 3a) such as Epcam, Sox9, Slug, Snail, Stat3, Klf4 or the recently identified Mmp3-Wnt5b axis (Fig. 3b). To further substantiate this finding, qPCR analysis showed higher CD133 and Nanog levels (Fig. 3c), and immunofluorescence staining revealed higher numbers of CD133-positive cells35 at the border of ADMs in AKC mice at 10 weeks of age (Fig. 3d,e). In line, Cxcr4, a marker labelling invasive cancer stem cells36, is also increased in AKC mice (Fig. 3f). Further evidence is provided by a hyperactive NODAL/SMAD2 signalling axis in AKC mice identified by both GSEA36 and immunoblotting for Nodal and phosphorylated SMAD2/3 (Supplementary Fig. 4f,g,h). Of note, a recent study shows that the overexpression of Nodal induces a metastatic phenotype in pancreatic cancer cells via the Smad2/3 pathway37.

Acinar cell de-differentiation via enhanced BMP4 signalling. EMT is associated with secretion and accumulation of growth factors that have a profound influence on the tumour micro-environment and support tumour growth26. Numerous signalling pathways, including BMPs and Nodal38,40, are known to converge in early development to direct EMT and cell lineage commitment41. GSEA of differentially regulated genes between AKC and KC mice identified a hyperactive BMP4 pathway in AKC mice (Fig. 4a, Supplementary Fig. 3c)42. Interestingly, we identified an enrichment of both BMP4 signalling and EMT factors and EMT-signalling gene sets in AKC mice at the 5 weeks

Figure 2 | Epithelial-to-mesenchymal transition (EMT) is accelerated in ATM-targeted pancreata. (a) Genome-wide transcriptional profiling identified 2,472 differentially regulated genes shown as a hierarchical clustered heat map of pancreata from 10 week aged p48fl/fl, KrasG12D/+;Atm−/− (AKC) - and p48cre/+;KrasG12D/+;Atm−/+ (KC) mice. (b,c) Immunohistochemistry (left panel, scale bar, 20 μm, right panel, scale bar, 10 μm) and quantitative analysis (c) of Sox9 positivity in precursor lesions. Representative images of normal acini of the respective genotypes are also shown to illustrate ductal differentiation. (d) GSEA of differentially regulated genes from (a) shows enrichment of EMT-associated genes in p48cre/+;KrasG12D/+;Atm−/+ pancreata using two independent ‘GSEA-core’ sets26,29. (e) RT-qPCR showing increased levels of Fibronectin, Twist1, Fsp1 and Vimentin in the pancreas of p48Cre/+;KrasG12D/+;Atm−/+ and p48Cre+/+;KrasG12D/+;Atm−/− versus controls (n = 5 per group). Student’s t-test *P<0.05. Error bars, s.e.m. (f) Immunohistochemistry staining shows expression of Vimentin and Fibronectin in precursor lesions. Scale bar, 20 μm. (g) Immunofluorescence stainings reveal more abundant expression of Zeb1-positive cells in CK19-positive precursor lesions of AKC-pancreata compared with controls. Scale bar, 20 μm. (h) Quantitative analysis of vimentin and Zeb1-positive cells in precursor lesions in the, respective, genotypes. (i) Immunofluorescence staining in p48cre/+;KrasG12D/+;Atm−/+ - and Rosa_tdxRFPfl/fl mice against RFP (red) and ZEB1 (green) and Dapi (blue). Scale bar, 10 μm. (j) RT-qPCR analysis showing expression levels of Sox9 (ADM: n = 4 versus 6; PanIN: n = 4 versus 5), Slug (ADM: n = 4 versus 6; PanIN: n = 4 versus 6) and N-cadherin (ADM: n = 3 versus 6; PanIN: n = 3 versus 6) in microdissected ADM or PanIN lesions from p48cre/+;KrasG12D/+;Atm−/+ and p48cre/+;KrasG12D/+;Atm−/− animals, respectively. Mann-Whitney test was used for statistical analysis. *P<0.05, **P<0.01. Error bars, s.e.m.
Figure 3 | ATM depletion enriches for cancer stem cells and associated signalling pathways. (a) Gene set enrichment analysis of differentially regulated genes from (Fig. 2a) identifies enrichment of a stem cell associated gene set in AKC pancreata at 10 weeks of age. (b) Hierarchically clustered heat map illustration shows differential expression of stemness-associated genes among AKC mice. (c) RT-qPCR showing increased levels of CD133 and Nanog in the pancreas of AKC mice versus controls. (d) Immunofluorescence staining of pancreata from the respective genotypes at 10 weeks old shows expression of CK19 (red), CD133 (green) and nuclei (Dapi-blue). Scale bar, 20 μm. (e,f) IHC staining and quantifications for CD133 (scale bar, 20 μm) (e) and Cxcr4 (scale bar, 10 μm) (f) in the respective genotypes.
**Figure 4 | Loss of ATM activity compromises acinar cell integrity.**

(a) GSEA of the differentially regulated genes from the respective genotypes identifies enrichment of the BMP4 signalling signature in p48Cre/+; KrasG12D/+; Atm−/− pancreata at 5 weeks and 10 weeks of age. (b) Immunoblot of BMP4, Phospho-Smad 1/5/8 and β-actin in the respective genotypes. (c) Quantification of several immunoblots as representative images in b for the following mouse numbers. Bmp4: 4 KC versus 6 AKC animals. SPmad 1/5/8: 4 KC versus 5 AKC animals. Mann–Whitney test was used for statistical analysis. *P < 0.05, **P < 0.01. Error bars, s.e.m. (d) Bmp4 staining of KC and AKC mice pancreata shows predominant Bmp4 staining in the acinar compartment of AKC mice. Staining is representative for at least three mice per group. Scale bar, 10 μm. (e) Bright-field images of acinar cell cultures from freshly isolated acini cultured for 2 days in growth factor reduced matrigel under indicated conditions. BMP4 was used at 25 ng ml−1; n = 4. Scale bar, 50 μm. (f,g) Quantification of ductal structures at day 2 of culture and RT-qPCR showing levels of the ductal gene marker CK19 in the respective conditions. Fold changes were calculated by setting levels in control treated Atm+/+ acini to 1. Mann–Whitney test was used for statistical analysis. *P < 0.05, **P < 0.01. Error bars, s.e.m. (h) High-power bright-field images of acinar cell cultures under the indicated conditions at day 2 of culture. Scale bar, 10 μm. (i) Quantification of ductal structures at day 2 and (j) RT-qPCR analysis showing levels of the ductal marker gene CK19 in the respective conditions from Atm−/− acini (2 out of 3 experiments with similar results are shown). Error bars, s.e.m. All analyses were performed on 4- to 6-week-old animals.
stage, underpinning the contribution of these factors to disease initiation and progression (Fig. 4a, Supplementary Fig. 3f). This was further substantiated by western blot analysis with significantly increased BMP4 precursor protein and phosphorylated SMAD1/5/8-levels observed in the pancreas of AKC mice (Fig. 4b,c). IHC analysis spatially localized the increased BMP4 signal in AKC mice to the acinar compartment, thereby pointing towards the acini as the primary source for local BMP4 activity in AKC mice (Fig. 4d). Recently, the BMP4/SMAD1/MMP2 axis has been identified to mediate EMT and subsequent invasive properties in PDAC\(^9\). Indeed, MMP2 is upregulated in AKC mice as confirmed via microarray and immunoblotting from whole pancreatic lysate (Supplementary Fig 3a,g).

To directly assess the propensity of ATM-deficient acini to trans-differentiate to duct-like cells, acinar cell clusters were explanted and grown in matrigel layers \textit{in vitro}. Of note, these experiments were performed in ATM-targeted mice and control mice in the absence of oncogenic Kras to more accurately define the role of ATM loss as a single event. Isolation of the acinar cell compartment was achieved with \textgreek{~}90\% cell viability (data not shown) from both control and AC mice. At day 1 of culture, no significant differences in acinar cell morphology were observed (Supplementary Fig. 5a). Remarkably, however, acini from AC mice underwent significantly more ductal reprogramming compared with controls by day 2 (Fig. 4e-g). Thus, ATM loss as a single event facilitates ductal programming. On the basis of our observation that BMP signalling was deregulated in AKC mice, we included BMP4 and inhibitors of the BMP4 signalling pathway in the acinar trans-differentiation process (Fig. 4h-j, Supplementary Table 1). Hierarchical clustering using Euclidean distance revealed closer association of AKC mice pancreata with the QM human PDACs compared with KC mice (Fig. 5e). We also found an independent human gene expression signature of 36 genes indicating poor prognosis in PDAC patients strongly enriched in 10-week-old AKC-mice\(^9\) (Fig. 5f). A recent study performed a mouse to human search for proteomic changes associated with pancreatic tumour development and identified a novel set of genes that robustly discriminated against pancreatic cancer cases from matched controls\(^1\). We found that this gene set was also over-represented in AKC versus KC mice (Fig. 5f). Collectively, these data suggest that our AKC mouse cohort displays a remarkable number of similarities to the molecular changes previously observed in more aggressive subtypes of human PDAC.

Next we investigated ATM expression using human PDAC tissue microarrays. Analysis of normal pancreatic tissues showed 92\% expressed abundant ATM protein levels (\(n = 12/13\)). In 57 PDACs, ATM levels in the neoplastic compartment were high in only 33\% (19/57) and low in 66\% (38/57) (Fig. 6a,b). Moreover, we identified an inverse correlation between ATM protein expression and World Health Organization (WHO) tumour grading. While high-ATM expression levels were detected in 44\% (\(n = 15/34\)) of low-grade carcinomas (WHO Grade I and II), this was the case in just 17\% (\(n = 4/23\)) of high-grade tumours (WHO Grade III and IV). This indicates that loss of ATM is associated with a less differentiated tumour phenotype. In addition, ATM-low tumours showed significantly more lymph node metastasis (Supplementary Table 1). Notably, ATM protein expression has previously been shown to correlate significantly with human pancreatic tumour invasion in separate cohorts\(^17,52\), independently complementing our findings.

To determine whether the results obtained from our mouse model were translatable to the human disease, we examined ATM expression levels as a predictor for on-going EMT in human PDAC. Indeed, we found that ATM-low cases had more isolated CK-positive invasive cells. These EMT-like features were present in 74\%, whereas the ATM-high cases showed such features in only 40\% (Fig. 6c,d). This indicates that loss of ATM expression is associated with a more infiltrative/EMT-like phenotype in human PDAC. On the basis of these findings, we examined ATM mRNA levels in a large cohort of surgically resected primary human PDACs. Intriguingly we found that patients with low levels of ATM (\(n = 39\)) had a significantly reduced survival rate compared with patients with high-ATM levels (\(n = 18\)) (Fig. 6e). Finally, we applied the BMP4 signalling signature used in our mouse model to stratify pancreatic cancer patients, which have been separated according to their ATM expression levels (high vs low). In line with our mouse model, GSEA identified significant enrichment of this signature in the ATM low expressing cohort. Of note, the stem cell signature was also enriched. Thus, we found that low ATM expression is associated with perturbed Bmp4 signalling and this is correlated with shortened survival in human PDAC patients (Fig. 6e-g).

In summary, our results indicate that deregulation of ATM is a contributing factor supporting PDAC initiation/progression.
As shown in Fig. 7, we identified that reduced levels of ATM coupled with oncogenic Kras activation resulted in a higher number of dysplastic pancreatic lesions. This was primarily due to loss of acinar cell identity and a gain in duct-like cell features (acinar-to-ductal metaplasia, ADM), followed by acinar-to-ductal reprogramming (ADR) and PanIN formation, which was broadly associated with an altered TGFβ superfamily signalling and EMT. In addition, these ductal precursor lesions were associated with a pronounced fibrotic reaction. Taken together, our data suggest an intimate link between ATM expression and PDAC progression in mice and men.

Discussion
ATM is a large serine/threonine kinase implicated in a plethora of cellular functions\(^5\). Current data suggest that the role of ATM in tumorigenesis is context and tissue dependent. Certain cell types become more sensitive to DNA-double strand break (DSB)-inducing agents upon ATM loss\(^5\). In contrast, the DNA-damage response acts as an inducible barrier in human gliomas frequently harbouring inactivating mutations in the Atm/Chk2/Trp53 axis\(^5\). In the current study, we note that neither members of the DDR nor p21 nor the apoptosis or senescence programmes are remarkably altered in response to...
loss of ATM. A recent study shows in lung and breast cancer cells that ATM acts as a binary switch to control the contribution of p53 signalling to the DNA damage response and to determine treatment response. Thus, further studies are warranted to investigate the Atm-Trp53 axis in pancreatic cancer in more detail. ATM has previously been shown to prevent dysplastic...
growth in the colon, independent of its effects on genomic stability\textsuperscript{56}. Thus, our data are in line with accumulating evidence that ATM integrates with a variety of signalling cascades and future studies will help unravel the role(s) of ATM apart from its established role in the DDR programme\textsuperscript{57}.

ADM in the pancreas is an emerging field of research and recently key transcription factors implicated in this process have been characterized. One such factor, SOX9, accelerates formation of pre-malignant lesions when concomitantly expressed with oncogenic K-RAS\textsuperscript{59}. Together with the EMT-promoting factor Slug, SOX9 determines the mammary stem cell state but also drives EMT. In turn, their co-expression increases tumorigenic and metastatic features of human breast cancer cells and is associated with poor patient survival\textsuperscript{56,57}. We identify that ATM loss in a mouse model of PDAC enhances SOX9 expression even in intact acinar structures, implicating an intimate connection between ATM and ductal programming. Interestingly, this link seems to be conserved among several cancers as ATM\textsuperscript{−/−} lymphoblastoid cells also show significantly increased SOX9 levels compared with their ATM\textsuperscript{+/+} counterparts\textsuperscript{57}. Ductal programming in ATM-depleted pancreas displayed close correlation with on-going EMT, shown by high levels of ZEB1 and other EMT markers in precursor lesions. Importantly we validated this independently of our microarray analysis by microdissection of neoplastic lesions and sorting for DBA+ cells. Given the large amount of neoplastic lesions in the pancreas at this time point, this is undoubtedly contributing to the microarray data in addition to the fibrosis. Notably, the EMT-TF ZEB1 is a marker for metastatic properties and stemness but also serves as an independent predictor of mortality in PDAC\textsuperscript{2}.

ATM loss was recently reported to enhance breast cancer stem cell properties via a TGFB\textsuperscript{−} dependent mechanism\textsuperscript{58,59}. In the current study, we found that loss of ATM alters TGFB-\beta superfamily signalling as shown by a perturbed BMP4/Smad1/5/8 and Nodal/Smad2/3 signalling axis. Nodal governs EMT via induction of Snail, and related paracrine signalling events help to establish a niche for tumour initiating cells as shown for pancreatic stellate cells\textsuperscript{39,60}. Moreover, high Nodal/Smad2/3 signalling levels in human PDAC cohorts determine poor survival\textsuperscript{61} and an EMT phenotype\textsuperscript{37}. Similarly, the BMP4/Smad1/MMP2 axis drives EMT in PDAC\textsuperscript{40} and notably, our data un.ravel a series of concomitant events aligned with these published observations. Intriguingly, loss of ATM activity in the pancreas gives rise to elevated BMP4 signalling that acts as a switch in the maintenance of acinar cell integrity and induces trans-differentiation to metastatic ductal cells. The phenotype observed fits well with previous models in which mice with intact Smad4 signalling display a TGFB\textsuperscript{−} independent, EMT-associated tumour growth\textsuperscript{62}.

Recent exome-sequencing data from human PDAC and respective progenitor lesions have positioned ATM among the 16 most commonly mutated genes in human PDAC\textsuperscript{5,15,16}. Most recently, ATM mutations were observed predominantly in genetically instable human PDACs\textsuperscript{2} and ATM was one of the few genes already mutated in the germline of PDAC patients\textsuperscript{32}. The current study provides molecular insight into the consequences of ATM mutations in PDAC. We highlight key aspects of ATM loss in murine tumour biology and demonstrate that this model faithfully recapitulates subtypes of human PDAC. Low-ATM levels are associated with increased BMP4 levels and elevated stem cell gene signatures, which have been previously linked to disease outcome in a variety of cancers including colorectal cancer\textsuperscript{33}. This supports the hypothesis that loss of ATM activity gives rise to more aggressive, EMT-rich tumours due to increased tumour initiating cell potential. Our data uncover that loss of heterozygosity of ATM in AKC mice is sufficient to reduce survival. This observation is in line with the previously reported higher risk of cancer risk and mortality in human ATM heterozygous-mutated patients\textsuperscript{29} and supports previous studies in mice\textsuperscript{22,27}. Also, sequencing analysis of human PDAC cases reported heterozygous mutations in the ATM gene\textsuperscript{2} and linked ATM to BRCA in genetically unstable PDAC subtypes\textsuperscript{3}. In line, ATM heterozygosity cooperates with loss of Brca1 to generate anaplastic breast cancer\textsuperscript{21,63}. It will be interesting to investigate how co-deletion of such factors affects PDAC formation. Taken together our current finding that loss of ATM is implicated in the early stages and progression of PDAC suggests new avenues of signalling mechanisms in PDAC. Therefore, this study facilitates considering ATM as a potential clinical target in human PDAC.

### Methods

**Mouse strains.** To generate p48Cre;Kras\textsuperscript{G12D+/−};Atm\textsuperscript{lox/lox} mice, we backcrossed the p48Cre and Kras\textsuperscript{G12D+/−} lines\textsuperscript{64,65} to the ATM\textsuperscript{lox/lox} line\textsuperscript{66} twice to generate p48Cre; Atm\textsuperscript{lox/lox} and Kras\textsuperscript{G12D+/−};Atm\textsuperscript{lox/lox} mice. We then crossed Kras\textsuperscript{G12D+/−} and ATM\textsuperscript{lox/lox} mice to p48Cre+/−;Atm\textsuperscript{lox/lox} mice to produce experimental animals on a mixed C57/BL/6, 129/BALB/c background. Representative genotyping is shown in Supplementary Fig. 1a. Some mice were also bred with Z/AP, a double reporter mice line as previously described\textsuperscript{67} or Rosa26 tdRFP/lox/lox mice to allow for tracing Cre recombinase expression driven by the p48 promoter (Supplementary Fig. 1b, 4b). Tail-derived DNA was used to confirm the genotype of mice from the breeding crosses using specific primers and representative genotyping is shown in Supplementary Fig. 1a. Murine genotypes followed expected mendelian frequency and primers used for genotyping are outlined in Table 2. All studies were performed under ethical and animal protection regulations of the University of Ulm.

**Histology and Immunohistochemistry.** Tissue specimens were fixed in 4% phosphate-buffered paraformaldehyde overnight at 4°C and embedded in paraffin. Immunohistochemical analyses were performed on serial sections of 4 μm using standard techniques. Antibodies used are outlined in the supplementary Materials and Methods. All images were taken with Olympus BX40 with spot Insight QE camera or Mirax Scan (Carl Zeiss), and the IF images were taken with a Zeiss EL-Einsatz Axioskop (Carl Zeiss).

**RNA isolation and Quantitative RT-PCR.** RNA was extracted according to the manufacturer’s instructions using the RNaseasy Mini or Micro Kits (Qiagen) and eluted in 40 or 15 μl RNAase-free H\textsubscript{2}O. cDNA was subsequently synthesized using the iScript cDNA synthesis kit (BioRad). Quantitative real-time RT-PCR analysis (qPCR) was carried out according to the manufacturer’s instructions. The PCR reaction was performed using the SensiMixSYBR kit (Bioline) in a Rotor-Genie 6000 series thermal cycler (Qiagen) using the following PCR reaction: denaturation at 95°C for 2 min, followed by 40 cycles at 95°C for 15 s, 60°C for 1 min and 72°C for 15 s. To verify the specificity of the PCR amplification products, melting curve analysis was performed. mRNA levels were normalized to Hmbs, cyclophilin A or Gapdh mRNA levels. Primer information is shown in Table 2. Primer sequences are not available for commercial primers from Qiagen.

**Histopathological analysis and scoring of pancreatic lesions.** Tumours were scored in blinded manner by a board-certified pathologist (J.K.L.). On the basis of a combination of H&E, Alcian Blue and CK19 stainings, ADM and PanIN lesions were
human PDA in Fig. 5e and minimize the batch effect, all samples (PDA samples) in NCBI GEO. To compare ATM samples and
and customized R script of GSEA. Human PDA tumour samples
classified according to histopathologic criteria. At least 10 fields at high-power magnification were imaged and characterized for
ARTICLE
differentiation all tumours were classified by using WHO ADM lesions and PanIn stage. For grading of tumour

at high-power magnification were imaged and characterized for
the genes that have fold change
probe sets. The differential expression analysis of any two conditions
R/Bioconductor. The expression l
scanned using a microarray scanner (Agilent Technologies). Raw
nologies) according to the manufac
was labelled with the Low Input Quick Amp Kit (Agilent Tech-
Kit (Agilent Technologies). F
as described above. Genome-wide gene expression profiles of indi-
individual KC and AKC mouse pancreata at 5 (AKC) or 10 (n = 3 KC; n = 4 AKC) weeks old was extracted as described above. Genome-wide gene expression profiles of indi-
ual AKC and KC samples were generated using the SurePrint G3 Mouse GE 8 × 60 K (Design ID 028005) Microarray
Kit (Agilent Technologies). Fifty nanograms of each sample was labelled with the Low Input Quick Amp Kit (Agilent Tech-
ologies) according to the manufacturer’s instructions. Slides were scanned using a microarray scanner (Agilent Technologies). Raw
data were pre-processed and quantile normalized using R/Bioconductor. The expression level of individual gene was then
calculated by averaging the signal intensities of all corresponding
probe sets. The differential expression analysis of any two conditions
was performed using limma t-test in R/Bioconductor. Accordingly, the genes that have fold change >1.5 and corrected P value < 0.05 were considered as being differentially expressed. Array data are available in GEO under accession code GSE68808. To gain functional
insight of differentially expressed genes, the enrichment analysis was conducted using pre-defined gene sets and customized R script of GSEA. Human PDA tumour samples were retrieved from GSE17891 (39 samples) and GSE17891 (26 PDA samples) in NCBI GEO. To compare ATM samples and human PDA in Fig. 5e and minimize the batch effect, all samples were subjected to COMBAT algorithm, resulting in a merged data set including 73 samples and 14,468 genes. Hierarchical clustering was performed in R. In Fig. 6e–g, pancreatic cancer data from ICGC were retrieved from GSE36924. The samples were divided into two, ATM-high (>7.6) and ATM-low (= < 7.6), groups on the basis of expression of ATM gene. Accordingly, the GSEA was performed on the data set to evaluate the significance of pre-defined gene sets.

Gene expression microarrays. In brief, total RNA extracted from
individual KC and AKC mouse pancreata at 5 (n = 3 KC; n = 3
AKC) or 10 (n = 3 KC; n = 4 AKC) weeks old was extracted as described above. Genome-wide gene expression profiles of indi-
and customized R script of GSEA. Human PDA tumour samples were retrieved from GSE17891 (39 samples) and GSE17891 (26 PDA samples) in NCBI GEO. To compare ATM samples and human PDA in Fig. 5e and minimize the batch effect, all samples were subjected to COMBAT algorithm, resulting in a merged data

set including 73 samples and 14,468 genes. Hierarchical clustering was performed in R. In Fig. 6e–g, pancreatic cancer data from ICGC were retrieved from GSE36924. The samples were divided into two, ATM-high (>7.6) and ATM-low (= < 7.6), groups on the basis of expression of ATM gene. Accordingly, the GSEA was performed on the data set to evaluate the significance of pre-defined gene sets.

Tissue microarrays. Commercially available tissue microarray
sections of human PDACs (A207IV and A207V AccuMax Array) were purchased from ISU ABXIS (Seoul, Korea) and used according to the manufacturer’s instructions (The human bio-
logical product provided was obtained legally, in compliance with applicable national and local laws, regulations and guidelines.). In brief, after deparaffinization, antigen retrieval was performed in Sodium Citrate solution (pH 6). Sections were incubated with rabbit anti-ATM (#sc-7230, Santa Cruz at 1:1,000 dilution) or

Western blot analysis. Immunoblotting was performed according to standard procedures. In brief, a piece of pancreas was frozen in liquid nitrogen immediately after killing the mice.

| Table 2 | Primers used for genotyping and qPCR analysis. |
|---|---|---|---|
| **Genotyping** | | |
| Cre001 | Biomers | 5’-accaggcgctatacaactg-t3` | 5’-ctacattggtccagccac-t3` |
| Cre002 | Biomers | 5’-ccctttacaacgcagctaggta-t3` | 5’-agctacggccatggctagta-t3` |
| Kras005 | Biomers | 5’-actccctaatgtctccctccgtcc-t3` | 5’-gcctactccgtccaccaataatctgc-t3` |
| Kras006 | Biomers | 5’-acatctagttatagttgaggtg-t3` | 5’-ggctacagaggttggcaca-t3` |
| ATM gf86723 | Biomers | 5’-tccttttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| ATM BAC13 | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| ATM BAC 7 | Biomers | 5’-aaaaatttttttttttttttt-t3` | 5’-ggctacagaggttggcaca-t3` |
| RosaRFPI (SH57A) | Biomers | 5’-tactagtctatagttgaggtg-t3` | 5’-ggctacagaggttggcaca-t3` |
| RosaRFPI (SH65A) | Biomers | 5’-tactagtctatagttgaggtg-t3` | 5’-ggctacagaggttggcaca-t3` |
| RosaRFPI (LO86) | Biomers | 5’-tactagtctatagttgaggtg-t3` | 5’-ggctacagaggttggcaca-t3` |
| **qPCR** | | |
| Amylase | Biomers | 5’-cagagacatgggtcagcaggt-t3` | 5’-acgttaaagctcacaacag-t3` |
| CK19 | Biomers | 5’-gagggttccagtaaattgg-t3` | 5’-gaggagaggctagcaagcc-t3` |
| Sox9 | Biomers | 5’-tccttttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| Cyclophilin A | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| Vimentin | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| alpha SMA | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| Slug | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| DBA | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| FSP1 | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| N-cadherin | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| Fibronectin | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| **Microdissection qPCR** | | |
| CK19 | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| Sox9 | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| N-cadherin | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
| Slug | Biomers | 5’-aatccctttcctttggctttc-t3` | 5’-acgttcagaggccagt-g3` |
Protein lysates were prepared using protein extraction buffer (4% SDS, 100 mM Tris-HCl) containing protease inhibitors and 1 mM phenylmethylsulfonyl fluoride (PMSF), and cell debris was removed by centrifugation at 4°C for 10 min at 14,000 r.p.m. Protein content was measured using a colorimetric assay (Bradford Biorad Assay, Biorad). Lysates (30 μg) were resolved by SDS–PAGE and transferred to a PVDF membrane (#IPVH00010, Immobilon-P Membrane, Millipore). Immunoreactive bands were visualized using chemiluminescence (Thermo scientific, Waltham, MA, USA) (Supplementary Fig. 6). Images were processed and analysed using the ImageJ software (http://rsbweb.nih.gov/ij/). Antibodies used are as follows: goat anti-Bmp4 (#sc-6896 Santa Cruz); rabbit anti-Nodal (#39953, Abcam) rabbit anti-Phospho Smad 1/5/8 (#9511, Cell Signaling); rabbit anti-Phospho Smad 2/3 (#3101, Cell Signaling); rabbit anti-ATM (#ab78, Abcam) all 1:1000, O.N. at 4°C and mouse anti-β-actin (#3101, Sigma) 1:50,000 for 1 h at RT.

Microdissection of ADM and PanIN lesions. Paraffin-embedded mouse tissues were sequentially sectioned in three to four 10-μm-thick sections and a single 3-μm-thick section and were mounted onto glass slides. Sections were then deparaffinized in xylene, rehydrated using graded ethanol and washed with distilled water. The 3-μm-thick section was stained in haematoxylin and eosin and coaxed with glass slides to properly assess the morphology of the progenitor lesions. The 10-μm-thick slides were stained in eosin only and were then transferred to manual microdissection. Within each slide ADM and PanIN lesions were dissected separately. RNA isolation was performed utilizing the Qiagen miRNeasy FFPE Kit following the manufacturer’s instruction, and cDNA synthesis was done using 150-200 ng of RNA. The samples were subjected to qRT-PCR as described above. EMT marker expression was normalized to Ck19 expression levels, and Mann–Whitney test was performed to determine statistical significance.

Acinar cell isolation and culture. Acinar cell explants were isolated and cultured as previously described with slight modification. In brief, the pancreas of mice aged 4-6 weeks was removed and placed in cold Hank’s Balanced Salt Solution (HBSS). The pancreas were then chopped into small pieces and then transferred to Collagenase P solution (0.2 mg ml⁻¹) for 10 min at 37°C. The digested pancreas was washed twice with HBSS containing 5% FCS and filtered through a 100-μm cell strainer (Greiner bio-one). The tissue suspension was then layered gently on a 30% FCS/HBSS solution and centrifuged at 4°C to pellet the acini. The cell pellet was then resuspended in a 1:2 media: growth factor reduced matrigel Beacton Dickinson (BD) solution and plated on a 24-well-cell culture plate. Media used for culturing the cells was Waymouth’s, 10% FCS, 1% P/S, 0.5% BSA, 5% horse serum, 1% penicillin/streptomycin, 1% insulin/transferrin/selenium (ITS), and 1% sodium pyruvate.
1 mg ml⁻¹ dexamethasone (Sigma) and 100 mg ml⁻¹ Soybean trypsin inhibitor (Sigma). Media was changed daily and human recombinant BMP4 (Preprotech 120-05) and human recombinant Noggin (Preprotech 120-10C) were used at 25 ng ml⁻¹ and 5 µM, respectively. Dorsomorphin and LDN-193189 (Sigma) were used at concentrations as indicated in Fig. 4h–j and Supplementary Fig. 5 (ref. 43). RNA was isolated at day 2 and PCR analysis was performed as described above. Quantification of ductal structures was performed by counting at least six individual fields at ×10 magnification in triplicate.

Statistical analysis. Contingency graph statistics were calculated using the Fisher exact test. All other tests for significance, unless otherwise stated, were performed using an unpaired Student’s t-test. Kaplan–Meier curves were calculated using the survival time for each mouse from all littermate groups. The log-rank (Mantel-Cox) test was used to test for significance differences between the groups. Statistical analysis was performed using GraphPad Prism 5 (GraphPad Software). Error bars represent the s.e.m.

Proliferation index. The proliferation index for ADMs and mPanINs was determined by counting the Ki67-positive nuclei (actively proliferating cells) per field for 10 fields at ×40 magnification, with at least four mice per group. Each field selected contained ADMs and mPanINs exclusively and if stroma cells were within the field, these were excluded from the count. Differences between groups were evaluated using the Student’s t-test, where P < 0.05 was considered to be significant.

Fibrotic content. Sections were stained with Masson–Goldner stain and histologic fibrosis was evaluated in at least four mice per group. Images of each section were recorded at ×2.5 magnification and fibrotic regions were defined as distinct blue/grey areas. The total area of each fibrotic region was measured and expressed as a percentage of the total section area. All measurements were recorded and analysed using Image J software. Differences between groups were evaluated using the Student’s t-test, where P < 0.05 was considered to be significant.

Senescence-associated-β-galactosidase (SA-β-gal) assay. Frozen sections of pancreatic tissue were fixed with 4% paraformaldehyde in PBS for 15 min, washed with PBS and stained at 37°C in the dark for 14–16 h in X-Gal solution (1 mg ml⁻¹ X-Gal, 40 mM Citric-acid Sodiumphosphate, 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide and 1 mM MgCl₂ in PBS at pH 5.3) and subsequently counterstained with Hoechst.

Cytokine antibody array. Proteins present within pancreata form the respective genotypes were screened using the RayBio Mouse Cytokine Antibody Array C series 1000 (RayBiotech) according to the manufacturer’s instructions. Total protein was isolated from pancreas using 1% Triton X-100 lysis buffer and centrifuged at 10,000 g for 5 min at 4°C. The crude supernatants were transferred to nitrocellulose membranes (Millipore) and blocked with 5% non-fat milk in TBS-T for 1 h. The membranes were then incubated with primary antibodies. After washing, the membranes were incubated with secondary horseradish peroxidase-conjugated antibodies. The membranes were then washed and developed with X-ray film (Amersham). The membrane was then imaged using a phosphor imager (BioRad). The protein expression levels were quantified using ImageJ software and normalized to the total protein levels.

Co-immunofluorescence for RFP and Zeb1. Pancreatic specimens were fixed in 4% phosphate-buffered paraformaldehyde overnight at 4°C and embedded in paraffin. Immunohistochemical analyses were performed on serial sections of 4 µm using standard techniques. In brief, after deparaffination, antigen retrieval was performed in Citrate buffer (pH 6). Slides were blocked with CAS-Block histochemical reagent (Invitrogen) for 45 min at RT. Anti-Zeb1 primary antibody was used at 1:100 dilution in blocking solution overnight at 4°C. Sections were then washed and incubated with secondary antibodies for 45 min at room temperature. Subsequently, biotinylated anti-RFP primary antibody diluted 1:200 in blocking solution was added to the sections for 1 h at RT. After washing Vectastain ABC-AP complex (Vector laboratories) was directly applied for 30 min at RT and developed using fluorescence Vector Red Alkaline Phosphatase (AP) Substrate Kit (Vector Laboratories). Controls using each antibody combination alone on serial sections were performed. All IF images were taken using a Zeiss EL- Einsatx Axioskop (Carl Zeiss).

ATM mRNA expression in human PDAC. In addition to newly generated data, we curated the freely available data portal of the International Cancer Genome Consortium (ICGC; www.icgc.org; last accessed 20 September 2013). In brief, we extracted patient characteristics, clinicopathological data (stage, grade and so on) and outcome data for survival analysis. The associated gene expression data (GSE36924) were extracted from GEO (http://www.ncbi.nlm.nih.gov/geo/; last accessed 20 September 2013). We analysed the frequency distribution of values from the ATM probeset (ILMN_1716231) using a non-linear, fourth polynomial fit. The local maximum ( > 7.6) was defined as the ATM cutoff for distinction of high versus low ATM.

Pancreatic ductal cell isolation. Pancreatic ductal cells were isolated according to a recently published protocol31. In brief, the pancreas of mice aged 9–12 weeks were removed and placed in G-solution (Hank’s Balanced Salt Solution (HBSS), 0.9 g l⁻¹ glucose and 47.6 µM CaCl₂). After washing, the pancreas were minced into small pieces ( < 1 mm²) using surgical scissors and scalpels. The tissue pieces were then transferred to DMEM/F12 containing 1 mg ml⁻¹ collagenase V (Sigma) and 100 mg ml⁻¹ Soybean trypsin inhibitor (Sigma) and incubated at 37°C for ~ 35 min with rotation. The reaction was stopped by adding cold G-solution, and the cell suspension was then centrifuged at 300 g for 5 min at 4°C. The cellular pellet was re-suspended in 2 ml trypsin-EDTA for 2 min at RT using a 1 ml pipette. The reaction was stopped and the cells were pelleted via centrifugation as above. The cellular pellet was subsequently washed with cold separation buffer (PBS, 0.5% BSA and 2 mM EDTA), filtered through a cell strainer and re-centrifuged. The cellular pellet was then separated into 400 µl aliquots and one was kept as a pre-sorting fraction. The other aliquots were subject to staining with DBA lectin-FITC (Vector Laboratories) at a dilution of 1:400 for 10 min on a rotor in the dark at 4°C. Following washing in the same buffer, cells were centrifuged at 300 g for 10 min. The cellular pellet was then resuspended in 90 µl separation buffer and 10 µl of anti-FITC Microbeads (Miltenyi Biotec) was added and the solution was incubated in the dark for 15 min at 4°C. After a final washing step, separation was performed with MS columns (Miltenyi Biotec), according to the manufacturer’s protocol. RNA from all fractions (DBA-positive, DBA-negative and presorting) was isolated using the RNasy microkit (Qiagen) and subject to cDNA synthesis as described above. Efficient separation of the different fractions was confirmed using DBA-specific primers (data not shown).
References


Acknowledgements

We thank U. Möhnle, S. Fischer, R. Köhntop and V. Sakk for excellent technical assistance. We also thank Milena Armacki for helpful discussions and the animal facility staff for mouse colony maintenance. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, KL 2544/1-1), Boehringer-Ingelheim BIU (N5 to S.L.; C1 to A.K.), Else-Kröner-Fresenius-Stiftung (2011_A206 to A.K. and S.L.), BMBF (Systar to A.K.). A.K. is indebted to the Baden-Württemberg Stiftung for the financial support of this research project by the Eliteprogramme for Postdocs. A.K. is an Else-Kröner-Fresenius Memorial Fellow. R.R. was supported by the Graduate College 1041, Ulm University.

Author contributions

This study was conceived by M.W., A.K., R.R., S.L.; experiments were designed by R.R., L.P., S.L., M.W., A.K.; H&E-stained and antibody stained sections were reviewed by R.R., L.P., A.K. and by M.W., who has extensive knowledge of similar mouse models of pancreatic cancer. Acquisition of all the data was done by R.R., L.P., A.L., Q.L., F.M.F., E.H., J.G., M.G., D.H., G.V.F., S.E.W.; the analysis and interpretation of data was performed by M.W., A.K., R.R., L.P., S.L., J.K.L., M.Z., K.-L.R.; the final manuscript was prepared by M.W., A.K., R.R., L.P., S.L.; the whole study was supervised by M.W., A.K., S.L., R.R.

Additional information

Accession codes: Array data are available in GEO under accession code GSE68808

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/


This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
Supplementary Figure 1. Characterization of the conditional ATM mouse model co-expressing oncogenic K-RAS. (A) Genotyping PCR from tail DNA shows expression of ATM wildtype and floxed alleles, p48Cre recombinase and Kras^{G12D} with the respective negative controls. (B) Representative image of the pancreas from Z/AP-reporter mouse in which p48Cre- activation results in a homogenous AP (in red) reporter signal indicating almost complete recombination within the exocrine compartment of the mouse pancreas. Blue staining delineates non-recombined islets. (C,D) Representative immunoblot and respective quantification of ATM expression in the pancreas of mice from the indicated genotypes. (E)
Representative images of the pancreata from p48Cre;Atm−/− mice pancreas over 1 year of age reveal normal pancreatic architecture.
Supplementary Figure 2. Immunohistochemical analysis of KC- and AKC mice for different markers. (A) Immunofluorescence staining of pancreata from the respective genotypes at 10 weeks old shows expression of CPA (red), Sox9 (green) and Nuclei (Dapi-blue) in precursor lesions. Scale bar, 20 μm. (B-G) Immunohistochemical staining reveals (B) Cyclin E (Scale bar, 10 μm) and Cyclin D (Scale bar, 20 μm), (C) p53 and p21 (Scale bar, 10 μm), (D) γH2AX (Scale bar, 10 μm), (E) SA-βgal (Scale bar, 10 μm; PDA image Scale bar
10 μm), (F) Tunnel and (G) CD45 expression in the respective genotypes and conditions (Scale bar, 20 μm). Irradiated and non-irradiated small intestines from mouse serve as positive and negative controls for γH2AX staining. Normal adult pancreas and PDAC serve as negative controls for SA-βgal. Representative images from at least 3 mice per group are shown. *P<0.05, **P<0.01, ***P<0.0001 (Student's t-test). Error bars, s.e.m.
Supplementary Figure 3. Heat map analysis and GSEA of microarray data on AKC and KC mice (A,B) Hierarchically clustered heat map illustration a differential expression pattern for numerous genes from the (A) MMP and (B) collagen families in pancreata from 10 week AKC- vs. KC-mice. (C) Hierarchically clustered heat map illustration shows differential expression of TGF-β superfamily members among AKC- and KC-mice. (D) A representative GSEA from 10 random gene sets each containing 100 genes randomly selected from the mouse genome shown in (E), shows no enrichment to the differentially regulated gene list shown in Figure 2A. All the p-values are above 0.05, suggesting no similarity or bias to any group of samples. (F) Gene set enrichment analysis of differentially regulated genes from Figure 2A identifies enrichment of the (E) EMT core gene set and slug based EMT gene set (gene sets taken from 2,3) in AKC - pancreata at 5 weeks of age. (G) A mouse cytokine array using total pancreatic lysate from a p48Cre;KrasG12D+/++;Atm+/+ (left) and a p48Cre+/+,KrasG12D+/++;Atm+/− (right) mouse. Error bars, s.e.m.
Supplementary Figure 4. ATM loss drives EMT and stemness (A,B) Illustration of strategy to generate p48Cre;Kras^{G12D/+};Atm^{-/-};Rosa_tdRFP^{fl/fl} mice (p48^{Cre/+} = “C”; Kras^{G12D/+} = “K”; Atm^{-/-} = “A”; Rosa_tdRFP^{fl/fl} = “R”) and tdRFP expression in recombined adult pancreas with
no expression found in the spleen. (C) Illustrative overview of the DBA-lectin sorting strategy. (D) RT-qPCR expression for Ck19 and amylase in the indicated cellular compartment following DBA-lectin MACS sorting. (E) RT-qPCR for Sox9, twist1, vimentin, Slug and N-cadherin in the DBA+ fractions from p48Cre+/;KrasG12D+/;Atm+/− and p48Cre;KrasG12D+/;Atm+/+ mice (n=3 per genotype). Error bars, s.e.m. (F) Gene set enrichment analysis of differentially regulated genes (from (Figure 2A) using previously described gene sets 4 identifies enrichment of the Nodal signalling pathway in AKC pancreata. (I,J) Immunoblot and quantification of Nodal, Phospho-Smad 2/3 and β-actin in the respective genotypes. Error bars, s.e.m.
Supplementary Figure 5

Supplementary Figure 5. Loss of ATM increases ADM formation. (A) Brightfield images of freshly isolated acinar cell cultured at 1 day from the indicated genotypes. Scale bar, 100 μm (B) RT-qPCR showing levels of the acinar differentiation marker – Amylase in the respective cell culture conditions at day 2. Error bars, s.e.m. (C) Low power brightfield (10 x magnification) images of acinar cell cultures under the indicated conditions at day 2 of culture. Scale bar, 50 μm.
Supplementary Figure 6

(A) Western blot analysis of BMP4 precursor in whole pancreas from various groups.

(B) Western blot analysis of p-Smad1/5/8 in whole pancreas from various groups.

(C) Western blot analysis of Nodal in whole pancreas from various groups.

(D) Western blot analysis of p-Smad2/3 in whole pancreas from various groups.

(E) Western blot analysis of β-actin in whole pancreas from various groups.

Supplementary Figure 6. Original gel pictures for western blots.

The figure shows uncropped western blots displayed in Fig. 4B and Supplementary Figure 4G. (A,B,C) Western blot analysis of protein levels of BMP4, phosphorylated-Smad1/5/8 and β-actin in whole pancreatic lysates from the indicated groups. (D,E) Western blot analysis of protein levels of Nodal and phosphorylated-Smad2/3 in whole pancreatic lysates from the indicated groups.
Supplementary Table 1. Pancreatic Cancer Cohort (commercial TMA)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All N=57</th>
<th>ATM-high* N=19</th>
<th>ATM-low* N=38</th>
<th>P pos vs. neg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td>0.39</td>
</tr>
<tr>
<td>Male</td>
<td>35</td>
<td>10</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>22</td>
<td>9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>61 (32-80)</td>
<td>63 (45-78)</td>
<td>61 (32-80)</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td>0.56</td>
</tr>
<tr>
<td>Head</td>
<td>38</td>
<td>12</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Body</td>
<td>13</td>
<td>3</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Tail</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>≤2cm</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>&gt;2cm</td>
<td>51</td>
<td>17</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>T-Category</td>
<td></td>
<td></td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td>T1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>47</td>
<td>16</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>N-Category</td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>N0</td>
<td>25</td>
<td>13</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>N1+</td>
<td>32</td>
<td>6</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Stage grouping</td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>I</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>27</td>
<td>6</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td></td>
<td></td>
<td></td>
<td>0.046</td>
</tr>
<tr>
<td>Grade 1</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Grade 2</td>
<td>28</td>
<td>13</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Grade 3</td>
<td>22</td>
<td>16</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Note: TNM stages and stage grouping follows AJCC 7th ed; P values from student’s t-test for age, Fisher’s exact test for dichotomous variables, or chi-square when taking all categories into account.

* A case with nuclear ATM staining was scored as “high” when there was more than 10% labelling within the tumour cell fraction and “low” when less than 10% were positive.
Supplementary References


9. Statutory Declaration

I hereby declare that I wrote the present dissertation with the topic: “Molecular insights into early cell fate specification and pancreatic cancer” independently and that I used no other aids than those cited. In each individual case, I have clearly identified the source of the passages that are taken word for word or paraphrased from other works.

I also hereby declare that I have carried out my scientific work according to the principles of good scientific practice in accordance with the current “Satzung der Universität Ulm zur Sicherung guter wissenschaftlicher Praxis” [Rules of the University of Ulm for Assuring Good Scientific Practice].

Ulm, den 27.07.15, Ronan Russell
10. Acknowledgements

First and foremost, I would like to thank PD. Dr. Alexander Kleger for giving me the opportunity to perform this work in his laboratory. His outstanding supervision, suggestions and support are reflected throughout the work presented here. Moreover, his support and drive have provided me with invaluable experience and to develop my career as a research scientist.

Prof. Dr. Michael Kühl has been fantastic in his support and encouragement during my thesis. His expertise in Biochemistry and Molecular Biology are greatly appreciated and provided a source of support throughout.

Prof. Dr. Martin Wagner has provided a wealth of knowledge particularly in supporting in vivo studies. I will be forever grateful for his patience and teaching skills.

I would also like to thank all members of the thesis committee for taking the time to review my thesis.

Ralf Köhntop who has always been a pleasure to work alongside. He provides a great backbone to the lab structure and always knows how to lighten the mood! Stefanie Fischer and Claudia Rhuland were always a fantastic support and more than willing to share their knowledge and experience. For this I will be forever grateful. Dr. Anett Illing and Meike Hohwieler and Dr. Clair Weidgang have been great colleagues and were always willing to share, discuss and troubleshoot scientific issues! In addition all the past and current members of the Kleger lab, the Seufferlein group, Uschi from Prof Wagner’s lab, the group of Stefan Liebau and all members of the Department of Internal Medicine I. You have provided support in many different ways and I am grateful to have had the pleasure to work alongside you all.

Finally, to my parents, family and friends. None of this would have been possible without your continuous support, generosity and understanding. For this I want to say a big thank you.
Name: Ronan (Michael) Russell

Nationality: Irish

Education:
April 2011 – July 2015 PhD student
Stem Cell Differentiation and Cell Reprogramming Lab - Dr. Alexander Kleger
Internal Medicine 1, University Hospital Ulm, Germany

October 2010 - April 2011 Internship, The Mucosal Immunity Laboratory
Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland

September 2009 - October 2010 Research MSc. Biochemistry and Molecular Biology
University College Dublin, Ireland

University College Cork, Ireland
Grade : Second Class Honours Grade 1
Thesis: The role of PKC isozymes in the developing mouse brain

September 2000 – May 2005 Leaving Certificate (Irish State Examination)
St. Joseph’s College, Thurles, Co. Tipperary, Ireland

Honours and Awards:
September 2014 Best Short Presentation, Viszeralmedizin 2014, Leipzig, Germany

June 2014 GSCN (German Stem Cell Network) Travel Award to attend Hydra X Summer School, September 2014

October 2013 Top Poster Prize, UEG Week, Berlin, Germany, 2013

April 2011 - April 2014 German Research Foundation / Deutsche Forschungsgemeinschaft - (DFG) Scholarship

September 2009 – June 2010 Swiss Government Scholarship for Postgraduate Research

Work experience:
June - September 2006 Quality Control, Transdermal pharmaceutical products
Alza Ireland (Johnson and Johnson Group), Cashel, Co. Tipperary, Ireland
Work experience:
June - September 2006 Quality Control, Transdermal pharmaceutical products
Alza Ireland (Johnson and Johnson Group), Cashel, Co. Tipperary, Ireland

Additional Information:
Certificate: RESAL Module 1 - Introductory Course in Laboratory in Animal Science (Switzerland)
Languages:
English: Mother tongue
French: Intermediate
German: Basic

Courses/Workshops:
September 2014 Hydra X Summer School on Stem Cell Biology
Hydra Island, Greece
*competition based selection to attend this course

July 2014
Hot topics in experimental GI cancer
United European Gastroenterology Basic Science Course
Klinikum Rechts der Isar, Technical University Munich, Germany
*competition based selection to attend this course

July 2012
Epithelial ion transport in the GI tract
United European Gastroenterology Basic Science Course
First Department of Medicine and Department of Pharmacology and Pharmacotherapy, University of Szeged, Hungary
*competition based selection to attend this course

July 2011
Bregenz Summer School on Endocrinology
European Society of Endocrinology (ESE) with the German Society of Endocrinology (DGE), Young Active Research in Endocrinology (YARE), Deutsche Gesellschaft für Osteologie (DGO)
Bregenz, Austria

June 2011
Murine Embryology: Embryo recovery and Immunostaining
Dr. Jennifer Nichols (PhD; FSB) Wellcome Trust Centre for Stem Cell Research
University of Cambridge, United Kingdom

Presentations:
September 2014 Title: Loss of ATM accelerates Pancreatic Cancer Formation and Epithelial - Mesenchymal Transition
Viszeralmedizin 2014 – Gastroenterologie – Viszeralchirurgie
Leipzig, Germany

September 2013
Title: Tbx3 regulates cell fate decisions during murine endodermal development
Viszeralmedizin 2013 – Gastroenterologie – Viszeralchirurgie
Nürnberg, Germany

July 2012
Title: Loss of ATM increases cancer formation in a mouse model of human
pancreatic ductal adenocarcinoma  
United European Gastroenterology Basic Science Course  
Szeged, Hungary

Conference Participation:

December 2014  
Cell Symposia: Stem Cell Energetics  
Poster Presentation: Loss of ATM accelerates pancreatic cancer formation and Epithelial-Mesenchymal-Transition.  
Berkeley - CA, USA

October 2013  
Cell Symposia: Using stem cells in modelling and treating disease  
Poster Presentation: Tbx3 directs cell fate decision towards mesendoderm.  
Cedars – Sinai Medical Center, Los Angeles, CA, USA

June 2013  
ISSCR 11th Annual Meeting  
Poster Presentation: Tbx3 directs cell fate decision towards mesendoderm.  
Boston, Massachusetts, USA

Publications:

*Faculty 1000 prime recommendation

2. TBX3 Directs Cell-Fate Decision toward Mesendoderm.  
#equal contribution

3. The Kvβ2 subunit of voltage gated potassium channels is interacting with ProSAP2/Shank3 in the PSD  
C. Proepper, S. Putz, R. Russell, T.M. Boeckers, S. Liebau  

4. Definitive endoderm formation from plucked-human hair derived induced pluripotent stem cells and SK channel regulation.  
doi:10.1155/2013/360573

5. Lipoxin A4 is a novel estrogen receptor modulator*  
Russell R. #, Gori I. #, Pellegrini C., Kumar R., Achtari C., Canny G.O.  
# equal contribution  
*Faculty 1000 prime recommendation
6. Tumor necrosis factor-α activates estrogen signaling pathways in endometrial epithelial cells via estrogen receptor α.
Gori I., Pellegrini C., Staedler D., Russell R., Jan C., Canny G.O.

Accepted manuscripts:

1. Loss of ATM accelerates pancreatic cancer formation and Epithelial-Mesenchymal Transition
Ronan Russell, Lukas Perkhofer, Stefan Liebau, Qiong Lin, Andre Lechel, Fenja M. Feld, Martin Zenke, Daniel Hartmann, Guido von Figura, Anett Illing, Stephanie E. Weissinger, Karl-Lenhard Rudolph, Peter Möller, Jochen K. Lennerz, Thomas Seufferlein, Martin Wagner, Alexander Kleger. Accepted at Nature Communications (May 2015)

Manuscripts in revision:

1. A dynamic role of Tbx3 in the pluripotency circuitry
Ronan Russell#, Marcus Ilg#, Qiong Lin#, Guangming Wu, Leonhard Linta, Meike Hohwieler, André Lechel, Wendy Bergmann, Pavan K Puvvula, Moritz Klingenstein, Olena Sakk, Stefanie Raab, Anne Moon, Martin Zenke, Thomas Seufferlein, Hans Schöler, Anett Illing, Stefan Liebau, Alexander Kleger
# equal contribution
In revision at Stem Cell Reports (April 2015)