The pioneer and differentiation factor FOXA2 is a key driver of yolk-sac tumour formation and a new biomarker for paediatric and adult yolk-sac tumours

Wasco Wruck1 | Felix Bremmer2 | Mara Kotthoff3 | Alexander Fichtner2 | Margaretha A. Skowron3 | Stefan Schönberger4 | Gabriele Calaminus5 | Christian Vokuhl6 | David Pfister7 | Axel Heidenreich7 | Peter Albers8 | James Adjaye1 | Daniel Nettersheim3

1Institute for Stem Cell Research and Regenerative Medicine, University Hospital Düsseldorf, Düsseldorf, Germany
2Institute of Pathology, University Medical Center Goettingen, Goettingen, Germany
3Department of Urology, University Hospital Düsseldorf, Düsseldorf, Germany
4Department of Pediatric Hematology and Oncology, University Children’s Hospital, Essen, Germany
5Department of Pediatric Hematology and Oncology, University Hospital Bonn, Bonn, Germany
6Institute of Pathology, University Hospital Bonn, Bonn, Germany
7Department of Urology, University Hospital Cologne, Cologne, Germany
8Department of Urology, University Hospital Düsseldorf, Düsseldorf, Germany

Abstract
Yolk-sac tumours (YSTs), a germ cell tumour subtype, occur in newborns and infants as well as in young adults of age 14-44 years. In clinics, adult patients with YSTs face a poor prognosis, as these tumours are often therapy-resistant and count for many germ cell tumour related deaths. So far, the molecular and (epi)genetic mechanisms that control development of YST are far from being understood. We deciphered the molecular and (epi)genetic mechanisms regulating YST formation by meta-analysing high-throughput data of gene and microRNA expression, DNA methylation and mutational burden. We validated our findings by qRT-PCR and immunohistochemical analyses of paediatric and adult YSTs. On a molecular level, paediatric and adult YSTs were nearly indistinguishable, but were considerably different from embryonal carcinomas, the stem cell precursor of YSTs. We identified FOXA2 as a putative key driver of YST formation, subsequently inducing \textit{AFP}, \textit{GPC3}, APOA1/APOB, ALB and GATA3/4/6 expression. In YSTs, WNT-, BMP- and MAPK signalling-related genes were up-regulated, while pluripotency- and (primordial) germ cell-associated genes were down-regulated. Expression of FOXA2 and related key factors seems to be regulated by DNA methylation, histone methylation / acetylation and microRNAs. Additionally,
INTRODUCTION

Testicular type II germ cell tumours (GCTs) represent the most common tumour of young men of age 17-45 years and incidence rates are rising steadily. Type II GCTs can be stratified into seminomas and non-seminomas, which both arise from the precursor lesion germ cell neoplasia in situ (GCNIS) as a result of a defective primordial germ cell (PGC) development. Seminomas are highly similar to GCNIS and PGCs with regard to morphology, gene expression and epigenetics. The non-seminomas have their own stem cell population—the embryonal carcinoma (EC). ECs are pluripotent and able to differentiate into cells of all three germ layers, resulting in formation of teratomas, and into extra-embryonic tissues, that is yolk-sac tumours (YSTs) and choriocarcinomas. In clinics, patients with YSTs face a poor prognosis, as YSTs count for many GCT-related deaths. YSTs frequently develop resistance towards the standard cisplatin-based therapy and cannot be cured by current standard treatment protocols. GCTs can also be found in newborns and infants, where these tumours are termed type I GCTs, which do not develop from GCNIS, but from an early defective PGC. They present mainly as teratomas and YSTs. Thus, in paediatric GCTs, occurrence and treatability of YSTs are also important issues.

The molecular and (epi)genetic mechanisms that control differentiation of ECs into YST are still unclear. In a previous study, we demonstrated that in vivo reprogramming of seminoma cells (TCam-2) into a non-seminoma-like cell fate (EC) can be induced by inhibiting the BMP-pathway. We identified SOX2, which was strongly up-regulated in response to BMP-pathway inhibition, as the key effector of this reprogramming process. In nude mice, TCam-2 cells deficient for SOX2 were not able to reprogramme to an EC anymore, but induced differentiation into YST-like tissues instead. Further analyses demonstrated that FOXA2, a pioneer and differentiation factor, might be the key driver of this YST-like differentiation, as FOXA2 interacts with typical YST-associated factors like AFP, APOA1, ALB and HAND1. Consequently, TCam-2 cells deficient for SOX2 and FOXA2 were not able to differentiate into non-seminoma-like cells at all. Thus, we suggest that FOXA2 might be the key effector in development of YSTs.

MATERIAL AND METHODS

2.1 Cell culture

All GCT cell lines were cultivated as described previously. See Table S1 for detailed information on cell lines (Table S1). STR profiles of all cell lines are checked on a regular basis and are available upon request.

2.2 RNA isolation

RNA was isolated from cell lines using the RNAeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol. RNA from frozen type I YST tissues was isolated by TRIzol reagent according to the manual (Qiagen).

2.3 Quantitative RT-PCR

cDNA synthesis and quantitative reverse transcription-polymerase chain reaction (qRT-PCR) were performed as described previously. Briefly, 1 µg of total RNA was in vitro transcribed into cDNA. Each sample was analysed in technical triplicates using 7.34 ng cDNA for each replicate. Oligonucleotide sequences are given in Table S2.

2.4 Immunohistochemistry

Immunohistochemistry (IHC) was performed as published previously. Briefly, antigen retrieval was carried out in citrate buffer. The primary antibodies were incubated for 30 minutes (min) at room temperature. Sections were incubated with a ready-to-use HRP-labelled secondary antibody at RT for 25 min. The substrate DAB + Chromogen system was used to visualize the antigen. Tissues were counterstained with Meyer’s haematoxylin. For antibody details, see Table S3.

2.5 Meta analyses of expression, DNA methylation and microRNA data

Affymetrix expression array raw data (CEL files) of paediatric and adult YSTs as well as ECs and human embryonic stem cells were read into the R/Bioconductor environment using the package ‘affy’. Data were transformed to the logarithmic (base 2) scale and normalized with the ‘Robust Multi-array average’ (RMA) method. Detection P-values were determined with the ‘MASS’ method implemented in the function ‘mas5calls’ from the package.
'affy'. Genes with detection P-values ≤ 0.05 were considered to be expressed. Pearson correlation coefficients between samples were calculated with the R-built-in function ‘cor’. Principal component analysis (PCA) of the samples was achieved through the R-built-in function ‘prcomp’, while for cluster analysis the R package ‘dendextend’ was employed. For identification of differentially expressed genes, linear models for microarrays as provided by the Bioconductor ‘limma’ package were fitted to the data. The p-values resulting from the ‘limma’ test were adjusted for the false discovery rate via the ‘qvalue’ package. Differentially expressed genes were assessed using the following criteria: detection P-value ≤ 0.05 at least in YSTs, limma-q-value ≤ 0.05, ratio ≤ 0.75 (down-regulation in YSTs); detection P-value ≤ 0.05 at least in ECs, limma-q-value ≤ 0.05, ratio ≥ 1.33 (up-regulation in YSTs); detection P-value ≤ 0.05 at least in ECs, limma-q-value ≤ 0.05, ratio ≤ 0.5 (down-regulation) and a limma-q-value ≤ 0.05.

2.5.2 | Meta-analysis—DNA methylation

Methylation data sets of human adult YSTs and ECs were downloaded from ‘The Cancer Genome Atlas’ (TCGA). Pre-processed DNA methylation data were used, and YSTs and EC samples were extracted from the testicular cancer data set. Beta-values provided in the pre-processed DNA methylation data were converted to M-values using the formula suggested by Du et al:

\[ M = \log_2 \left( \frac{\beta}{1 - \beta} \right) \]

Quality control procedures included Pearson correlation coefficients as described above and cluster dendograms via the R package ‘dendextend’. Differentially higher DNA methylation in YSTs was determined by the criteria: M ≥ 2 (in YSTs) or M ≤ -2 (in ECs), log\(_2\) ratio ≥ 2 (log ratio = M\(_{\text{YST}}\) - M\(_{\text{EC}}\)) and limma-q-value ≤ 0.05. Differentially lower DNA methylation in YSTs was determined by the criteria: M ≥ 2 (in ECs) or M ≤ -2 (in YSTs), log\(_2\) ratio ≤ -2 and limma-q-value ≤ 0.05.

2.5.3 | Meta-analysis—microRNAs

Analogously to DNA methylation data, microRNA data sets of human adult YSTs and ECs were downloaded from TCGA. Pre-processed microRNA data providing RPKM (reads per kilobase exon per million reads) values were used, and YST and EC samples were extracted from the TGCT data set. For determination of differential expression, a detection level of RPKM ≥ 1 (for up-regulation in YSTs, for down-regulation in ECs) was used together with a ratio ≥ 2 (up-regulation) or ratio ≤ 0.5 (down-regulation) and a limma-q-value ≤ 0.05. MicroRNA targets (homo sapiens) were downloaded from the TargetScan 7.2 database.

2.6 | Analysis tools

Venn diagrams were generated using Venny 2.1. Quality control procedures included Pearson correlation coefficients as described above and cluster dendograms via the R package ‘dendextend’ was employed.11 Differentially expressed genes were assessed using the following criteria: detection P-value ≤ 0.05 at least in YSTs, limma-q-value ≤ 0.05, ratio ≥ 1.33 (up-regulation in YSTs); detection P-value ≤ 0.05 at least in ECs, limma-q-value ≤ 0.05, ratio ≤ 0.5 (down-regulation) and a limma-q-value ≤ 0.05.

3 | RESULTS

To identify factors that are involved in YST formation, we meta-analysed gene expression microarray data of paediatric YSTs (pYST), adult YSTs (aYST) and ECs as well as human embryonic stem cells (hESCs) as controls (Data S1A–E). First, we compared the gene expression profiles of all samples to each other using a correlation matrix, unsupervised hierarchical clustering and a principle component analysis, demonstrating a high similarity of pYST to aYSTs (YST cluster) and among EC samples (EC cluster) (Figure S1A–C). The EC samples grouped with the hESCs and clearly apart from the p/aYSTs, while pYSTs and aYSTs were highly similar to each other (Figure S1A–C).

Next, we identified all genes differentially expressed between aYSTs and ECs (Data S1B). We found 126 individual genes up-regulated and 186 down-regulated in aYSTs compared with ECs (fold change (FC) >4) (Data S1C).

By using the STRING algorithm combined with a Gene Ontology (GO) search, we predicted interactions and the involved biological processes of up- and down-regulated genes (Figure 1A). Among the genes up-regulated in aYST versus ECs, we found FOXA2 and SOX17 as well as many FOXA2-related genes, like AFP, GPC3, APOA1/A2/B, ALB, TTR, FGA/B/G and DKK1,2,5,22 (Figure 1A, green labelled). Furthermore, interaction of GATA differentiation factors (GATA3/4/6; red labelled), WNT signalling-related (ANKRD6, BAMB1, BMP2, CDH1, DKK1, FRZB, FZD7, GATA3, GPC3, ISL1, LGR5, ROR2, SALL1, SFRP1, SOX17, TN1K, VANG1L1; blue labelled), BMP signalling-related (BAMBI, BMP2, CE1, DKK1, GATA3/4/6, GPC3, ROR2, SFRP1; turquoise labelled) and MAPK signalling-related factors (ACKR3, AGT, ANKRD6, BMP2, C5, CDH2, DKK1, DUSP4/9, ERBB4, FGA/B/G, FZD7, GPR37, KIT, NRG1, ROR2, SFRP1, TN1K, VEGFA; purple labelled) was predicted in aYSTs. Additionally, three ‘Cancer / Testis-Antigen’ (CTA) members of the MAGE family were up-regulated (MAGEA2/3/12; orange labelled).

In contrast, in aYSTs we found strong down-regulation of pluripotency / EC / PGC-related factors, such as NANOG, OCT3/4, KLF4, DPPA4, UTF1, FGFR, DNM7B, L1TD1, TFAP2C and DND1 (Figure 1B, red labelled). Additionally, metallothioneins (Figure 1B, green labelled), HLA molecules (HLA-DPA1/B, -DQA1/B1, -DRA

""
FIGURE 1  (A, B) STRING-based interaction prediction of genes up-regulated (A) or down-regulated (B) in aYST versus ECs. FOXA2 target genes are highlighted in green (A) according to the 'Harmonizome' database. Genes belonging to the same 'biological process' GO category were labelled by colours as indicated.
**FIGURE 2**  (A) Table summarizing numbers of analysed GCT tissues based on histology (left) and evaluation of IHC stainings (right). (B) HE staining and IHC for FOXA2, SOX17, GPC3, AFP and GATA3 in aYST tissues. (C) HE staining and FOXA2-IHC in pYST tissues. (D) OCT3/4- and FOXA2-IHC in mixed tumours with EC (white arrows) and aYST (black arrows) components. (E) HE staining and FOXA2-IHC in mixed tumours with teratoma (TER; white arrows) and aYST (black arrows) components. (F) HE staining and FOXA2-IHC in mixed tumours with choriocarcinoma (CC; white arrows) and aYST (black arrows) components. (G) HE staining and FOXA2-IHC in a seminoma (SEM) patient with elevated serum AFP levels. FOXA2-positive aYST cells could be found (black arrow). (H) qRT-PCR analysis of pluripotency and YST marker genes in pYSTs ($n = 6$), seminoma cells (TCam-2), EC cells (2102EP, NCCIT, NT2/D1), EC-YST cells (1411H) and aYST cells (GCT72). GAPDH and ACTB were used as housekeepers and for normalization.
-DNA) (Figure 1B, yellow labelled), CXCL and chemokine factors (CXCL9/10/11/12/13; CCR7, CCL5/8/19/21) (Figure 1B, blue labelled), ‘cluster of differentiation’ (CD) genes (CD2/3/D9/37/47/48/52/53/6/9-79A) (Figure 1B, pink labelled) and ‘interferon inducible proteins’ (IFN6/27/44/44L/H1/T1) (Figure 1B, patrol labelled) were down-regulated in aYST compared with ECs.

To find differences between pYSTs and aYSTs, we screened for differentially expressed genes (Data S1D) and identified only 19 genes (FC > 4; 12 up-regulated, 7 down-regulated in aYSTs vs. pYSTs) (Data S1E). Among them, MageA2/A3/A12 and cell cycle-related genes CCND2 (up-regulated in aYSTs) and CDKN1C (down-regulated in aYSTs) (Data S1E). Thus, pYST and aYST are highly similar with regard to gene expression.

We verified results by IHC on 342 FFPE-GCT-tissues (Figure 2A). We stained all samples for SALL4, OCT3/4, FOXA2, SOX17, GPC3, AFP and GATA3 (Figure 2A). All examined aYST populations (n = 117) showed a strong nuclear expression of FOXA2 (100%) and GATA3 (100%), while being focally positive for SOX17 (100%), GPC3 (93%) and AFP (95%) (Figure 2A,B). pYSTs were also positive for FOXA2 (100%) (Figure 2C). In mixed GCTs, FOXA2 clearly distinguished EC (Figure 2D, white arrows), teratoma (Figure 2E, white arrows) and choriocarcinoma (Figure 2F, white arrows) components from YST populations (Figure 2D-F, black arrows). All analysed tissues were positive for the GCT marker SALL4 (Figure 2A).29,30 We also analysed three seminoma patients showing increased levels of serum AFP. In these samples, FOXA2-positive cells could be detected, demonstrating presence of a YST component (Figure 2G, black arrow). Taken together, FOXA2 presents as a new and highly specific biomarker for p/aYSTs and is able to detect single YST cells in mixed GCTs. FOXA2 might also be a valuable biomarker to detect occult YSTs, e. g. in seminomas with elevated AFP levels.

By qRT-PCR, we analysed expression of pluripotency and YST marker genes in pYSTs (n = 6), the seminoma cell line TCam-2, three EC cell lines (2102EP, NCIT, NT2/D1), an EC-aYST cell line (1411H) and an aYST cell line (GCT72) (Figure 2H). 1411H growths as an EC in NT2/D1 EC and HepG2 hepatocellular carcinoma cells (Figure S3) as well as chemokine signalling (CCR7, CXCL9, CXCL12). Vice versa, Group 2 represents microRNAs down-regulated during YST formation and involved in repressing the pluripotency and PGC programme (KLF4, BCAT1, DND1, TFA2C) as well as chemokine signalling (CCR7, CXCL9, CXCL12). Group 2 represents microRNAs downregulated in YSTs and leading to de-repression of typical YST-associated genes, such as FOXA2, SOX17, AFP, APOA1, APOB, GATA3/4/6, BMP2, BAMBI, FZD7 and DUSP4.

Our results suggest that expression of FOXA2, which we postulate as a p/aYST key factor, is not regulated by DNA methylation, but might be repressed in ECs by microRNA1246 (Figure 3C). Furthermore, we asked, if FOXA2 expression might be regulated by epigenetic modifications on histone level. Thus, we screened ChiP-seq data extracted from the ENCODE project of various activating or repressing histone marks throughout the FOXA2 genomic locus in NT2/D1 EC and HepG2 hepatocellular carcinoma cells (Figure S3). HepG2 cells strongly express FOXA2 and many factors also found up-regulated during YST formation (AFP, APOA1/A2/B, ALB, FGA/B/G, GATA3/4/6, etc.; Figure S2B) and thus represent a valuable
FIGURE 3 (A) Genes deregulated (FC ≥ 4) in aYSTs versus ECs and showing inverse correlation to DNA methylation. (B) Deregulated microRNAs in aYSTs versus ECs. (C, D) Waterfall diagrams of expression dynamics (FC ≥ 4) of putative target genes of indicated microRNAs.
A}

early PGC

Microenvironment

FOXH2

SOX17+

pluripotency promoting

SOX17+

differentiation promoting

pYST

B}

EC

Microenvironment

FOXH2

SOX2+/SOX17-

pluripotency promoting

SOX2-/SOX17+

differentiation promoting

aYST

miR1246
H3K27me3

FOXH2

SOX17

miR1248
miR1293
miR1323

miR1246

WNT- / BMP- / MAPK-signaling

5mC

Pluripotency

Metallothioneins

Chemokine receptors / ligands

HLA molecules

CTA - MAGE

H3K4me1/3
H3K37me3
H3K9ac

FOXH2

SOX17

miR592
miR1263
miR1264
5mC

miR375

miR592

WNT- / BMP-signaling

Pluripotency

Metallothioneins

Chemokine receptors / ligands

miR592

HLA molecules

CTA - MAGE

induction

5mC DNA hypermethylation

repression

5mC DNA hypomethylation

inhibition
proxy for studying FOXA2 interactions. In FOXA2-NT2/D1 cells, we found high levels of repressive H3K27me3 mark around the FOXA2 transcription start site (red labelled) and throughout the gene body, while FOXA2+ HepG2 cells harboured high levels of activating H3K4me1/me3, H3K9ac and transcription promoting H3K36me3 (Figure S3). Thus, we propose that in ECs FOXA2 expression is silenced by miRNA1244 and repressive epigenetic modifications, for example H3K27me3, which need to be removed to allow for expression during differentiation. Additionally, induction of FOXA2 in p/aYSTs might be accompanied by activating histone modifications like H3K4me1/3, H3K9ac and H3K36me3.

4 | DISCUSSION

In this study, we deciphered the molecular and epigenetic mechanisms that regulate formation of YSTs and highlight FOXA2 as a new biomarker for p/aYSTs. We demonstrated that pYSTs and aYSTs are highly similar to each other with regard to gene expression, but aYSTs were clearly distinguishable from ECs based on gene and microRNA expression as well as DNA methylation. Thus, we assume that molecular mechanisms found in aYSTs are similarly detectable in pYSTs. Additionally, therapeutic options tested in aYSTs might also apply for pYSTs.

In our study, we highlight FOXA2 as a key driver of p/aYST formation acting in concert with SOX17, which already has been found up-regulated in p/aYSTs versus seminomas. FOXA2 is a pioneer and endodermal transcription factor expressed in several tumour types including genitourinary cancers, such as bladder carcinomas and prostate cancer and playing a crucial role in cellular differentiation. We suggested previously that during in vivo differentiation of TCam-2 cells into a non-seminoma including aYST-like structures, SOX17 switches function from pluripotency-promoting to differentiation-inducing as a result of FOXA2 up-regulation. Interestingly, partnering of SOX2 with PAX6 (instead of OCT3/4) leads to a switch in the function of SOX2 from pluripotency-promoting to endodermal differentiation-inducing. Given the high similarity between SOX factors and that SOX2 and SOX17 share similar functions in regulating pluripotency in ECs and seminomas, respectively, it seems reasonable that both factors are able to switch their functions in dependency of the interacting partner. The finding that SOX17 might be a key factor of YST formation strengthens the idea that pYSTs, which do not originate from a GCNIS, arise from an early SOX17+ PGC, where SOX17 switches function as a result of a microenvironment-triggered FOXA2 induction (Figure 4A). We suggest further that in SOX17/- SOX2+ ECs, a microenvironment-triggered FOXA2 induction (accompanied by inhibition of miR1246, removal of repressive H3K27me3 and establishment of activating H3K4me1/3, H3K9ac and H3K36me3) is an initial step in aYST formation, leading to up-regulation of SOX17 (Figure 4A). We propose that in ECs, seminomas and PGCs, SOX17 is able to switch from a pluripotency-promoting to a differentiation-inducing factor upon microenvironment-triggered FOXA2 induction, driving differentiation into YST lineage (Figure 4A).

Our data suggest that up-regulation of FOXA2 leads to induction of p/aYST-associated genes SOX17, AFP, APOA1/A2/B, ALB, TTR, FGA/B/G and GATA3/4/6 etc (Figure 4B). Interestingly, expression of these factors (including FOXA2) is highly typical for liver cells, that is hepatocytes. We also detected high expression of these factors in hepatocellular carcinomas (of both sexes) (Figure S2B). Furthermore, the WNT and BMP pathways play an important role in differentiation of human pluripotent stem cells to hepatocytes. Interaction of FOXA2 with many of the YST-related genes and signalling pathways has been shown in various settings including liver development and hepatocellular carcinomas. Additionally, in stem cell context SOX17 is an endodermal differentiation factor. Thus, on a molecular level, YST cells are closely related to liver (carcinoma) cells.

We detected up-regulation of WNT, BMP and MAPK signalling factors in aYSTs versus ECs. In parallel, the pluripotency programme was shut down in YSTs. In pYSTs versus germinomas, high activity of BMP and WNT signalling has already been shown. So, activation of WNT, BMP and MAPK signalling, while the pluripotency network becomes inactivated are important steps in p/aYST formation or maintenance.

Furthermore, metallothioneins, chemokine receptors / ligands and HLA molecules become down-regulated in YSTs versus ECs (Figure 4B). It remains elusive, if deregulations in these pathways are directly linked to activation of the FOXA2 axis or occur independently (Figure 4B).

We found that many of these key factors and driver processes might be controlled by DNA methylation and microRNAs (Figure 4B). During aYST formation, DNA methylation seems to be involved in silencing EC-associated genes, especially pluripotency and PGC genes, such as NANOG, BCAT1, DPPA4 and DND1 (Figure 4B). In line with this finding, we already demonstrated that NANOG expression is regulated by DNA methylation in GCTs, with NANOG-negative aYSTs harbouring with 70% the highest levels of NANOG promoter DNA methylation of all analysed GCT tissues (3.8% in seminomas, 6% in ECs, 66% teratomas and 62% choriocarcinomas). In this study, we highlight FOXA2 as a promising biomarker able to detect p/aYSTs with high specificity and to distinguish YST...
components from other GCT entities. Recently, another factor from the hepatocyte nuclear factor family (HNF1b) has been highlighted as a YST biomarker, further demonstrating the high similarity between YSTs and liver cells and suggesting that combining FOXA2 (HNF3b) and HNF1b to detect YSTs with high specificity in pathological routine diagnostics seems reasonable. Additionally, FOXA2 was sensitive enough to detect few YST cells in bulk tumour masses, like classical seminomas with elevated serum AFP levels, which has an important implication for clinical use; i. e. if FOXA2 is detectable in a patient's tumour sample, early more aggressive treatment may be recommended before YST outgrowth renders these progressive treatment-resistant GCTs incurable.

ACKNOWLEDGEMENT

We kindly thank Anna Pehlke for excellent technical assistance.

CONFLICT OF INTEREST

The authors confirm that there are neither conflicts of interest nor competing interests.

AUTHOR CONTRIBUTIONS

Wasco Wruck: Data curation (equal); Formal analysis (equal); Investigation (equal); Software (equal); Visualization (equal).

Felix Bremmer: Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing-review & editing (equal).

Mara Kotthoff: Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal).

Margaretha Skowron: Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (equal); Writing-review & editing (equal).

Alexander Fichtner: Formal analysis; Investigation; Methodology; Validation; Visualization.

Stefan Schönberger: Methodology; Resources (equal); Validation; Writing-review & editing.

Gabriele Calaminus: Resources (equal); Validation; Writing-review & editing.

Christian Vokuhl: Resources (equal); Validation; Writing-review & editing.

David Pfister: Resources (equal); Validation; Writing-review & editing.

Axel Heidenreich: Resources (equal); Validation; Writing-review & editing.

James Adjaye: Conceptualization (equal); Project administration (equal); Resources (equal); Supervision (equal).

Daniel Nettersheim: Conceptualization (lead); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (equal).

ETHICAL APPROVAL

The ethics committee (EC) of the Heinrich Heine University Düsseldorf (EC-HHU-D) raised no concerns about utilizing GCT cell lines for in vitro experiments (vote 2018-178). All type I GCTs were cryopreserved and anonymized during the paediatric GCT MAKEI 96-study and provided by the MAKEI 96-biobank with no concerns raised by EC-HHU-D about analysing these samples (votes 837 and 2019-822). Ethical approval for using the type II GCTs in the present study was obtained from the EC of the University Medical Centre Göttingen (vote 18/2/16).

DATA AVAILABILITY

Transcriptome data from YSTs and ECs were downloaded from the 'National Center for Biotechnology Information Gene expression omnibus' (NCBI GEO). In order to minimize technical variation, the meta-analysis comprises only data sets generated on the Affymetrix Human Genome U133A platform (GSE3218, GSE10783, GSE7332, GSE8481, GSE10615). Gene expression, MicroRNA, DNA methylation and mutational burden data were extracted from the TCGA cohorts 'Testicular GCT', 'Liver Hepatocellular Carcinoma' and 'Broad Cancer Cell Line Encyclopedia'. Histone-chromatin-immunoprecipitation-sequencing (Histone-ChIP-seq) data sets (ENCSR000EXA/B/C/D/E/F/EWZ; ENCSR000AMB/AMD/AME/AOL/APV/ATD, ENCSR755RRX) were extracted from the ‘Encyclopedia of DNA elements’ (ENCODE) project.

ORCID

Wasco Wruck https://orcid.org/0000-0002-9153-7051
Margaretha A. Skowron https://orcid.org/0000-0003-2152-384X
Stefan Schönberger https://orcid.org/0000-0002-7222-296X
Gabriele Calaminus https://orcid.org/0000-0002-7893-7982
Peter Albers https://orcid.org/0000-0002-1747-9615
Daniel Nettersheim https://orcid.org/0000-0002-4483-845X

REFERENCES

1. Cheng L, Albers P, Berney DM, et al. Testicular cancer. *Nat Rev Dis Prim*. 2018;4(1):29. https://doi.org/10.1038/s41572-018-0029-0
2. Park JS, Kim J, Elghiaty A, Ham WS. Recent global trends in testicular cancer incidence and mortality. *Medicine*. 2018;97(37):e12390. https://doi.org/10.1097/MD.0000000000012390
3. Berney DM, Looijenga LHJ, Idrees M, et al. Germ cell neoplasia in situ (GCNIS): evolution of the current nomenclature for testicular pre-invasive germ cell malignancy. *Histopathology*. 2016;69(1):7-10. https://doi.org/10.1111/his.12958
4. Nettersheim D, Looijenga LHJ, Idrees M, et al. Germ cell neoplasia in situ (GCNIS): evolution of the current nomenclature for testicular pre-invasive germ cell malignancy. *Histopathology*. 2016;69(1):7-10. https://doi.org/10.1111/his.12958
5. Nettersheim D, Gillis AJM, Looijenga LHJ, Schorle H. TGF-β1, EGF and FGF4 synergistically induce differentiation of the seminoma cell line TCam-2 into a cell type resembling mixed non-seminoma. *Int J Androl*. 2011;34(4pt2):e189-e203. https://doi.org/10.1111/j.1365-2605.2011.01172.x
6. Nettersheim D, Haimsoeth A, Jostes S, et al. SOX2 is essential for in vivo reprogramming of seminoma-like TCam-2 cells to an embryonal carcinoma-like fate. *Oncotarget*. 2016;7(30):47095-47110. https://doi.org/10.18632/oncotarget.9903
7. Nettersheim D, Vadder S, Jostes S, et al. SOX2 is essential for in vivo reprogramming of seminoma-like TCam-2 cells to an embryonal carcinoma-like fate. *Oncotarget*. 2016;7(30):47095-47110. https://doi.org/10.18632/oncotarget.9903
8. Skowron MA, Vermeulen M, Winkelhausen A, et al. CDK4/6 inhibition presents as a therapeutic option for paediatric and adult germ cell cancer. *PLoS One*. 2015;10(7):e013711. https://doi.org/10.1371/journal.pone.013711
cell tumours and induces cell cycle arrest and apoptosis via canonical and non-canonical mechanisms. Br J Cancer. 2020;123(3):378-391. https://doi.org/10.1038/s41416-020-0891-x

9. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004. https://doi.org/10.1186/gb-2004-5-10-r80

10. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy – Analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20(3):370-315. https://doi.org/10.1093/bioinformatics/btg405

11. Galili T. dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics. 2015;31(22):3718-3720. https://doi.org/10.1093/bioinformatics/btv428

12. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004;3(1):1-25. https://doi.org/10.2202/1544-6115.1027

13. Storey JD. A direct approach to false discovery rates. J R Stat Soc B Stat Methodol. 2002;64(3):479-498. https://doi.org/10.1111/1467-9868.00346

14. The Cancer Genome Atlas. Natl Cancer Inst Natl Hum Genome Res Inst. 2013.

15. Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinf. 2010;11(1). https://doi.org/10.1186/1471-2105-11-587

16. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective miRNA target sites in mammalian mRNAs. Elife. 2015. https://doi.org/10.7554/eLife.05005

17. Oliveros JC. An interactive tool for comparing lists with Venn Diagrams. http://bioinfgp.cnbc.es/tools/venny/index.html

18. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D613. https://doi.org/10.1039/nar-gk1y131

19. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cbioPortal. Sci Signal. 2013;6(269):pl1. https://doi.org/10.1126/scisignal.2004088

20. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012. https://doi.org/10.1038/nature11247

21. Davis CA, Hitz BC, Sloan CA, et al. The Encyclopedia of DNA elements (ENCODE): Data portal update. Nucieic Acids Res. 2018. https://doi.org/10.1039/nar-gk1x081

22. Kent WJ, Sugnet CW, Furey TS, et al. The Human Genome Browser at UCSC. Genome Res. 2002;12(6):996-1006. https://doi.org/10.1101/gr.229102

23. Rouillard AD, Gundersen GW, Fernandez NF, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford). 2016;2016:baw100. https://doi.org/10.1093/database/baw100

24. Korkola JE, Houldsworth J, Feldman DR, et al. Identification and validation of a gene expression signature that predicts outcome in adult men with germ cell tumors. J Clin Oncol. 2009;27(31):5240-5247. https://doi.org/10.1200/JCO.2008.20.0386

25. Korkola JE, Houldsworth J, Chadalavada RSV, et al. Downregulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors. Cancer Res. 2006;66(2):820-827. https://doi.org/10.1158/0008-5472.CAN-05-2445

26. Palmer RD, Barbosa-Morais NL, Gooding EL, et al. Pediatric malignant germ cell tumors show characteristic transcriptome profiles. Cancer Res. 2008;68(11):4239-4247. https://doi.org/10.1158/0008-5472.CAN-07-5560

27. Kami D, Shiojima I, Makino H, et al. Gremlin enhances the determined path to cardiomyogenesis. PLoS One. 2008;3(6):e2407. https://doi.org/10.1371/journal.pone.0002407

28. Barberi T, Bradbury M, Dincer Z, Panagiotakos G, Socci ND, Studer L. Derivation of engraftable skeletal myoblasts from human embryonic stem cells. Nat Med. 2007;13(5):642-648. https://doi.org/10.1038/nm1533

29. Cao D, Li J, Guo CC, Allan RW, Humphrey PA. SALL4 is a novel diagnostic marker for testicular germ cell tumors. Am J Surg Pathol. 2009;33(7):1065-1077. https://doi.org/10.1097/PAS.0b013e318a13eef

30. Miettinen M, Wang Z, McCue PA, et al. SALL4 expression in germ cell and non-germ cell tumors. Am J Surg Pathol. 2014;38(3):410-420. https://doi.org/10.1097/pas.0000000000000116

31. Vogelzang NJ, Bronson D, Savino D, Vessella RL, Fraley EF. A human embryonal–yolk sac carcinoma model system in athymic mice. Cancer. 1985;55(11):2584-2593.

32. Motallebipour M, Ameur A, Reddy Bysani MS, et al. Differential binding and co-binding pattern of FOXA1 and FOXA3 and their relation to H3K4me3 in HepG2 cells revealed by ChiP-seq. Genome Biol. 2009;10(11):R129. https://doi.org/10.1186/gb-2009-11-r129

33. Wang K, Brems JJ, Gamelli RL, Holterman AX. Foxa2 may modulate hepatic apoptosis through the ciAP1 pathway. Cell Signal. 2013;25(4):867-874. https://doi.org/10.1016/j.cellsig.2012.12.012

34. Palmer RD, Murray MJ, Saini HK, et al. Malignant germ cell tumors display common MicroRNA profiles resulting in global changes in expression of messenger RNA targets. Cancer Res. 2010;70(7):2911-2923. https://doi.org/10.1158/0008-5472.CAN-09-3301

35. Lee CS, Friedman JR, Fulter JT, Kaestner KH. The initiation of liver development is dependent on Foxa transcription factors. Nature. 2005;435(7044):794. https://doi.org/10.1038/nature03649

36. Yamashita H, Ampsona VO, Warrick JI, et al. On a FOX hunt: Functions of FOX transcriptional regulators in bladder cancer. Nat Rev Urol. 2017;14(2):98-106. https://doi.org/10.1038/nruro.2016.239

37. Park JW, Lee JK, Witte ON, Huang J. FOXA2 is a sensitive and specific marker for small cell neuroendocrine carcinoma of the prostate. Mod Pathol. 2017;30(9):1262-1272. https://doi.org/10.1038/modpathol.2017.44

38. Zhang S, Bell E, Zhi H, et al. OCT4 and PAX6 determine the dual function of SOX2 in human ESCs as a key pluripotent or neural factor. Stem Cell Res Ther. 2019;10(1). https://doi.org/10.1186/s13287-019-1228-7

39. Jostes SV, Fellermeyer M, Arévalo L, et al. Unique and redundant roles of SOX2 and SOX17 in regulating the germ cell tumor fate. Int J Cancer. 2020. https://doi.org/10.1002/ijc.32714

40. Irie N, Weinberger L, Tang WW, et al. SOX17 is a critical specifier of human primordial germ cell fate. Cell. 2015;160(1-2):253-268. https://doi.org/10.1016/j.cell.2014.12.013

41. de Jong J, Stoop H, Gillis AJ, et al. Differential expression of SOX17 and SOX2 in germ cells and stem cells has biological and clinical implications. J Pathol. 2008;215(1):21-30. https://doi.org/10.1002/path.2332

42. Ang LT, Tan AKY, Autio MI, et al. A roadmap for human liver differentiation from pluripotent stem cells. Cell Rep. 2018;22(8):P2190-2205. https://doi.org/10.1016/j.celrep.2018.01.087

43. Gosalia N, Yang R, Kerschner JL, Harris A. FOXA2 regulates a specific marker for small cell neuroendocrine carcinoma of the prostate. Mod Pathol. 2017;30(9):1262-1272. https://doi.org/10.1038/modpathol.2017.44

44. Sinner D. Sox17 and -catenin cooperate to regulate the transcription of endodermal genes. Development. 2004. https://doi.org/10.1242/dev.01176
45. Sinner D, Rankin S, Lee M, Zorn AM. Sox17 and β-catenin cooperate to regulate the transcription of endodermal genes. *Mol Cell Biol*. 2007; 27(22):7802–7815. https://doi.org/10.1242/dev.01176

46. Li Z, Tuteja G, Schug J, Kaestner KH. Foxa1 and Foxa2 are essential for sexual dimorphism in liver cancer. *Cell*. 2012. https://doi.org/10.1016/j.cell.2011.11.026

47. Fustino N, Rakheja D, Ateek CS, Neumann JC, Amatruda JF. Bone morphogenetic protein signalling activity distinguishes histological subsets of paediatric germ cell tumours. *Int J Androl*. 2011;34(4pt2):e218-e233. https://doi.org/10.1111/j.1365-2605.2011.01186.x

48. Joshi SS, Handorf E, Strauss D, et al. Large-scale sequencing of testicular germ cell tumour (TGCT) cases excludes major TGCT predisposition gene. *Eur Urol*. 2018; 73(6):978. https://doi.org/10.1016/j.eururo.2018.01.031

49. Skotheim RI, Lind GE, Monni O, et al. Differentiation of human embryonal carcinomas in vitro and in vivo reveals expression profiles relevant to normal development. *Cancer Res*. 2005;65(13):5588-5598. https://doi.org/10.1158/0008-5472.CAN-05-0153

50. Nettersheim D, Biermann K, Gillis AJM, Steger K, Looijenga LHJ, Schorle H. NANOG promoter methylation and expression correlation during normal and malignant human germ cell development. *Epigenetics*. 2011;6(1):114-122. https://doi.org/10.4161/epi.6.1.13433

51. Gallo A, Fankhauser C, Hermanns T, et al. HNF1β is a sensitive and specific novel marker for yolk sac tumor: a tissue microarray analysis of 601 testicular germ cell tumors. *Mod Pathol*. 2020;33(11):2354-2360. https://doi.org/10.1038/s41379-020-0597-x

52. Calaminus G, Goebel U, Teske C, et al. Prognosis and outcome of testicular germ cell tumors (GCTs) – Final analysis of the German mhao 94/98 and the MAKEI 98/96 trials. *Pediatr Blood Cancer*. 2010;55(5):806.

53. Edgar R, Edgar A. The Gene Expression Omnibus (GEO): A Gene Expression and Hybridization Repository. *NCBI Handb*. 2002.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.