INSTITUTE OF HUMAN GENETICS POLISH ACADEMY OF SCIENCES,

POZNAN, POLAND

A Dominant Gain-of-Function NANOS1 Variant Disrupts WNT Signalling and Impairs Human Germ Cell Specification

Doctoral thesis

by

Matisa Alla

Prepared under the supervision of

Prof. dr hab. Jadwiga Jaruzelska

This dissertation was prepared within Funding by National Science Centre Poland (OPUS 2019/35/B/NZ1/01665 to JJ).

Poznań 2025

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AmLC	Amnion-like Cell
BMP	Bone Morphogenetic Protein
Dox	Doxycycline
EB	Embryoid Body
eCLIP	Enhanced Cross-Linking Immunoprecipitation
ECM	Extracellular Matrix
EMT	Epithelial-to-Mesenchymal Transition
FACS	Fluorescent Activated Cell Sorting
GO	Gene Ontology
GOF	Gain of Function
GSEA	Gene Set Enrichment Analysis
hESC	human embryonic stem cell
hPGC	human primordial germ cell (*in vivo*)
hPGCLC	human primordial germ cell like cell (*in vitro*)
IF	Immunofluorescence
ME	Mesendoderm
Pre-me	Premesendoderm
RBP	RNA Binding Protein
RNP	Ribonucleoprotein
Tet-ON	Tetracycline Inducible System
TGF-ß	Transforming Growth Factor Beta
WNT	Wingless-related integration site
WRE	WNT-responsive enhancer elements

ABSTRACT

Infertility affects approximately 15% of couples worldwide, with genetic factors often underlying its etiology. NANOS1, an RNA-binding protein that regulates germ cell development post-transcriptionally through target mRNA 3′UTRs, has emerged as a critical fertility determinant from flies up to humans. A specific p.[Pro34Thr; Ser78del] variant (MUT-NANOS1) in NANOS1 is implicated in the absence of germ cells in the seminiferous tubules of infertile male patients, indicating a possible pathological mechanism. Preliminary studies in a seminoma-derived cell line showed that overexpression of MUT-NANOS1 substantially reduced cell viability, whereas wild-type (WT-) NANOS1 exerted anti-apoptotic effects.

To investigate how this variant influences early human primordial germ cell (hPGC) development, we employed a well-characterized hESC model offered by Prof Azim Surani from the University of Cambridge. Under defined in vitro conditions, transient mesendoderm-like precursors (pre-me) aggregate into embryoid bodies and differentiate into primordial germ cell–like cells (PGCLCs) at 10–40% efficiency under BMP stimulation. The W15 (46, XY) hESC line, equipped with a NANOS3-tdTomato fluorescent reporter, enabled real-time PGCLC tracking. We generated doxycycline-inducible WT- and MUT-NANOS1 lines and performed enhanced cross-linking immunoprecipitation (eCLIP) coupled with RNA-sequencing to identify bound and differentially expressed mRNAs.

Our data reveal that the p.[Pro34Thr; Ser78del] NANOS1 variant exerts a premature, repressive gain-of-function effect by downregulating canonical WNT signalling in pre-me cells, consequently impairing regulators such as EOMES and diminishing PGC competence. Notably, in PGCLCs, MUT-NANOS1 insufficiently represses WNT/TGF-β components, leading to aberrant WNT pathway activation, EMT-like transitions, and a shift toward mesendoderm/mesoderm fates. This disruption is further compounded by potential MUT-NANOS1 binding at the 5′UTR of NANOG and the 3′UTR of OCT4, suppressing these core pluripotency factors and reducing germ cell markers, including SOX17, TFAP2C, and NANOS3. Together, these alterations help explain the diminished *in vitro* PGC pool and gonadal germ cell depletion observed in patients carrying MUT-NANOS1. Notably, WNT inhibition partially restores pluripotency and germ cell markers while reducing BMP4 levels, highlighting the vital interplay among BMP, WNT, and TGF-β in safeguarding germ cell fate.

This study underscores how MUT-NANOS1 disturbs the transcriptome and signalling dynamics necessary for early germ cell specification, offering mechanistic insight into male infertility associated with NANOS1 p.[Pro34Thr; Ser78del] mutation.

STRESZCZENIE

Niepłodność dotyka około 15% par na świecie i często spowodowana jest czynnikami genetycznymi. NANOS1 to białko oddziałujące z 3'UTR cząsteczek mRNA, o kluczowej roli w potranskrypcyjnej regulacji rozwoju komórek gametogenicznych, od muszki owocowej po człowieka. Specyficzny wariant p.[Pro34Thr; Ser78del] (MUT-NANOS1) białka NANOS1 powiązany jest z brakiem komórek gametogenicznych w kanalikach plemnikotwórczych pacjentów, wskazując potencjalny mechanizm patologiczny. Wstępne badania nadekspresji MUT-NANOS1 wykonanej na linii ludzkich komórek gametogenicznych TCam-2 wyprowadzonej z ludzkiego nasieniaka pokazały istotne osłabienie żywotności komórek w porównaniu z komórkami z nadekspresją WT-NANOS1, w których zaobserwowano efekt anty-apoptotyczny.

Celem badań była próba wyjaśnienia jaki jest wpływ powyższego wariantu białka NANOS1 na wczesny etap rozwoju ludzkich komórek gametogenicznych (hPGC). W badaniach zastosowano wcześniej scharakteryzowany model linii ludzkich komórek hESC podarowany przez prof. Azima Suraniego z Uniwersytetu Cambridge. W zdefiniowanych warunkach in vitro, uzyskiwano najpierw przejściowy etap prekursorów mesendodermalnych (pre-me) stopniowo agregujących tworząc ciała embroidowe. W ich wnętrzu dochodziło do różnicowania komórek podobnych do pierwotnych komórek gametogenicznych (PGCLCs) pod wpływem stymulacji BMP z wydajnością 10-40%. Opisana wyżej linia ludzkich komórek embrionalnych (46,XY) W15 zaopatrzona we fluorescencyjny reporter NANOS3-dtTomato, umożliwiała obserwację powstawania PGCLCs w czasie rzeczywistym. Na podstawie powyższej linii komórkowej W15 wygenerowano linie komórkowe z nadekspresją MUT- oraz WT-NANOS1 indukowaną doxycykliną. Zastosowano technikę eCLIP celem identyfikacji mRNA z którymi każde z tych białek oddziałuje oraz RNAseq by uchwycić zmiany poziomu tych mRNA.

Badania pokazały represywny efekt „zysku funkcji" (gain of function) wariantu p.[Pro34Thr; Ser78del] przejawiający się m. in., osłabieniem kanonicznej sygnalizacji WNT w stadium pre-me. Konsekwencją było zaburzenie regulatorów różnicowania takich jak EOMES oraz obniżenie kompetencji komórek do różnicowania jako PGCLCs. Białko MUT-NANOS1 niewystarczająco obniżało komponenty ścieżki WNT/TGF-β, co prowadziło do nieprawidłowej aktywacji ścieżki WNT, podobnej do EMT tranzycji oraz przesunięcia różnicowania ku mesendodermalnemu/mezodermalnemu. Efekt ten był wsparty oddziaływaniem MUT-NANOS1 z 5'UTR mRNA NANOG oraz 3'UTR OCT4. Konsekwencją była supresja głównych czynników pluripotencji oraz redukcja markerów PGC takich jak, SOX17, TFAP2C oraz NANOS3. Powyższe zmiany pozwalają wyjaśnić obniżoną liczbę komórek PGCLC in vitro oraz brak w gonadach pacjentów z wariantem MUT-NANOS1. Należy podkreślić że zahamowanie ścieżki WNT skutkuje częściowym odtworzeniem pluripotencji oraz markerów PGC z jednoczesnym obniżeniem poziomu BMP4. Wskazuje to na istotne wzajemne oddziaływanie pomiędzy BMP, WNT oraz TGF-β ochraniające komórki gametogeniczne. Badania pokazują zaburzenie transkryptomu oraz dynamiki sygnalizacyjnej przez mutację p.[Pro34Thr; Ser78del], istotnych dla wczesnych etapów rozwoju komórek płciowych i pozwalają wniknąć w przyczynę niepłodności pacjentów obarczonych tą mutacją.

INTRODUCTION

1. Germline Associated Human Infertility

Infertility affects 10–15% of couples, with equal contribution from both partners, yet current treatments remain limited. Although techniques like IVF and ICSI can bypass infertility in many cases, true therapeutic solutions to restore germ cell numbers or regenerate reproductive organs require a deeper understanding of the biology of the testis, ovary, and germ cell development (Hayashi et al., 2012b). Moreover, male infertility, which increases the risk of testicular germ cell tumors up to 20-fold and by 52% among first-degree relatives (Rajpert-De Meyts et al., 2016; Nagirnaja et al., 2018), exhibits significant genetic overlap with female reproductive disorders and certain cancers, as evidenced by recent omics-based studies (Nagirnaja et al., 2018). Studies in mouse models (Cooke and Saunders, 2002) and human genetic analyses have further highlighted a significant role of genetic factors, including chromosomal abnormalities and single-gene mutations, in severe infertility cases (Hayashi et al., 2012b). Notably, mutations that disrupt proper germ cell specification and development can profoundly impair reproductive health, leading to long-term consequences for future generations (Castillo-Venzor et al., 2023).

2. Mammalian Primordial Germ Cells (PGCs)

The germ line is the sole heritable lineage in an organism, transmitting genetic information across generations (Oulhen et al., 2022). In mammals, PGCs, the precursors of the germ line, are specified early in embryonic development. Following their specification, PGCs migrate to the developing gonads, where they enter meiosis and differentiate into gametes. Fertilization of the oocyte by a sperm cell forms a totipotent zygote, capable of generating all cell lineages of the organism, including the germ line Thus, PGC specification represents a vital early step in establishing totipotency and ensuring the perpetuation of the mammalian life cycle (Tang et al., 2016). In humans, however, the precise timing of PGC specification is less clearly understood compared to mouse models, largely due to the practical difficulties with studying early post-implantation human embryos.

2.1 Differences in germ cell specification between mice and humans

Humans and mice diverged approximately 60 million years ago, resulting in distinct embryonic development structures: mice have acquired a unique egg cylinder formation during perigastrulation, while humans and other non-rodent mammals develop as a planar embryonic disc (Tang et al., 2016) (**Introduction Fig. 1**). In mice, PGC specification takes place in the posterior epiblast prior to gastrulation (**Introduction Fig. 1**). This process relies on the maintenance of pluripotency and complex signaling interactions between embryonic and extraembryonic tissues within the egg cylinder (Ohinata et al., 2009; Aramaki et al., 2013). These reciprocal signaling interactions between the epiblast, visceral endoderm (VE), and extraembryonic ectoderm (ExE) initiate gastrulation in the posterior epiblast at E6.5 (**Introduction Fig. 1**), leading to the formation of the

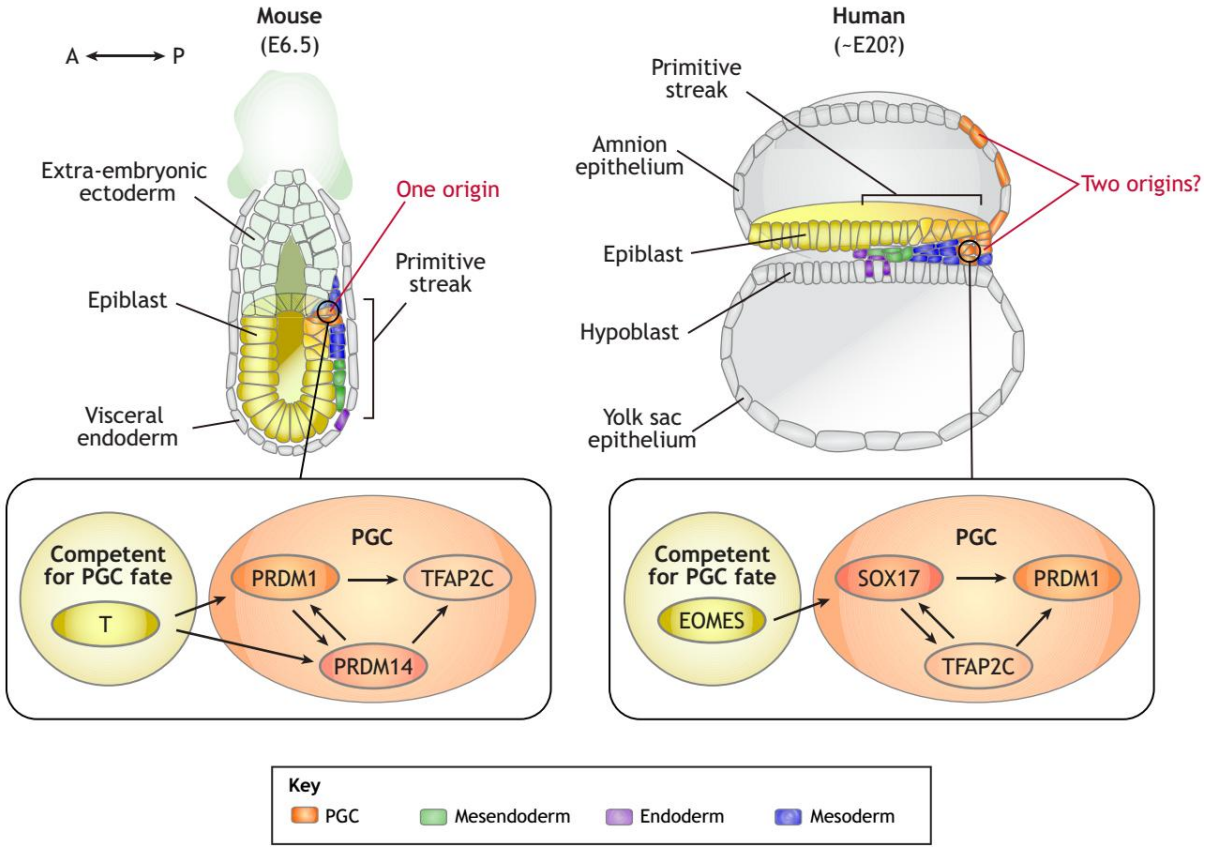
primitive streak (PS), which generates the mesoderm and definitive endoderm (Tam and Loebel, 2007), as well as considered to be the sole origin site of the PGCs in mice models (**Introduction Fig. 1**) (Kobayashi and Surani, 2018). Similarly, human hPGCs are thought to originate in the posterior epiblast of pre-gastrulation embryos. However, recent research on non-human primate embryos suggests that the nascent amnion (Sasaki et al., 2016), which arises from the post-implantation epiblast, may also serve as a site for PGC specification implying a potential two origin site for PGC specification (Kobayashi and Surani, 2018) (**Introduction Fig. 1**). In primates, including humans, the amnion forms from the pluripotent epiblast shortly after implantation, although the timing differs from species like pigs and other mammals, where the amnion develops later, following the onset of gastrulation (Hassoun et al., 2010).

Therefore, differences in PGC specification mechanisms are expected between mice and humans due to the diverse regulation of pluripotency and early post-implantation development in these species (Tang et al., 2016) (**Introduction Fig. 1**). Studies on *in vivo* migrating and gonadal PGCs in mice (mPGCs) and humans (hPGCs) has revealed a mix of conserved and species-specific features within the germ cell lineage (Tang et al., 2015). Conserved elements include germ cell specifiers such as BLIMP1 (PRDM1) and TFAP2C (AP2γ) **(Introduction Fig. 1)**, germ cell factors like NANOS3, DND1, DDX4, and DAZL, and pluripotency factors OCT4 (POU5F1) and NANOG. Notably, hPGC specifiers, including TFAP2C, cooperate to directly upregulate and maintain the core pluripotency factors OCT4 and NANOG. Before this upregulation, the pluripotency factors expression levels are lower, yet remain sustained because approximately 80% of the enhancers in transcription factors regulating OCT4 and NANOG are already poised or primed in human embryonic stem cells (hESCs) (Tang et al., 2022). However, while these factors are shared between mice and human embryos, their functional roles may differ significantly due to species-specific regulatory mechanisms (Kobayashi and Surani, 2018). Human PGCs exhibit several unique features not observed in their mouse counterparts.

For example, SOX17 is expressed in the human germline (**Introduction Fig. 1**) but not in the mouse germline, where SOX2 plays a pivotal role. Interestingly, while SOX2 is essential for maintaining mPGCs (Campolo et al., 2013), it is repressed in hPGCs (Irie et al., 2015; Tang et al., 2015). Similarly, KLF4, a factor associated with naïve pluripotency, is expressed in hPGCs but is repressed in mPGCs by PRDM1 (Durcova-Hills et al., 2008; Hackett et al., 2017).

Another key difference is the expression of EOMES in the human germline (**Introduction Fig. 1**), which acts upstream of SOX17 in PGC-competent epiblast cells (Kojima et al., 2017) and later in definitive endoderm (DE) (Teo et al., 2011). In contrast, brachyury (T, TBXT) (**Introduction Fig. 1**), is the equivalent mesodermal factor crucial for initiating mPGC fate (Aramaki et al., 2013) but appears to play a lesser role in hPGC specification compared to EOMES (Kojima et al., 2017). Notably, while PRDM1-positive mPGC precursors are generated in T-knockout mouse embryos, they are unable to maintain PRDM1 expression and fail to upregulate mPGC specifier PRDM14 (Tang et al., 2016), a transcription factor specifically linked to mPGC specification (**Introduction Fig. 1**). Therefore, while the conserved specifiers PRDM1 and TFAP2C

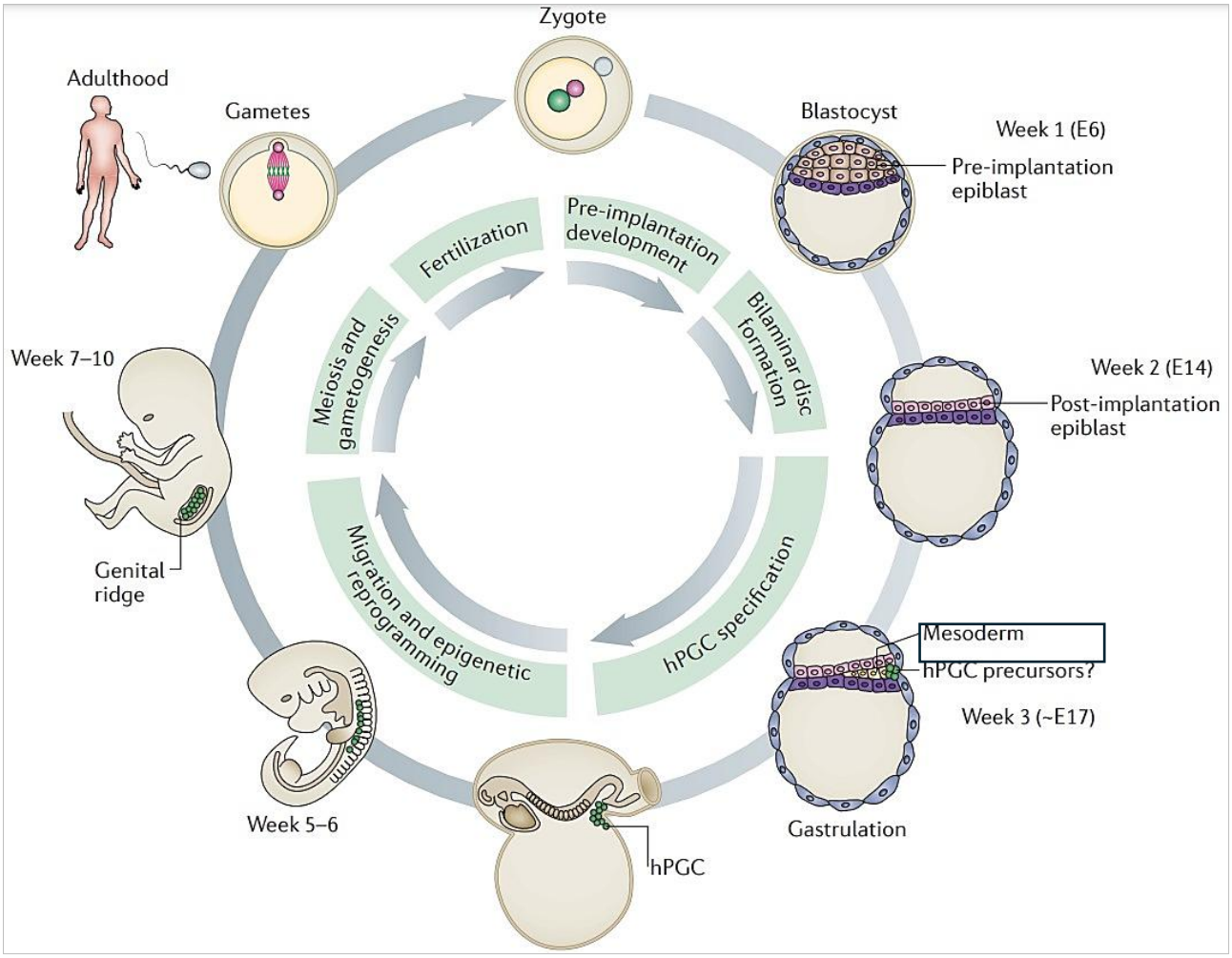
are critical for hPGC fate, their regulatory networks and downstream targets, as inferred from mutant cells and their probable targets analyses, differ between mice and humans (Irie et al., 2015; Kojima et al., 2017; Sasaki et al., 2015; Tang et al., 2015). For instance, bone morphogenetic protein (BMP) signalling is a conserved and shared pathway for germ cell induction in both humans and mice, but its activation and sustenance, shaped by epigenetic reprogramming, differs between human and mouse PGCs. Notably, in humans, SOX17 serves as the key epigenetic regulator for germline specification, whereas in mice, PRDM1 plays that critical role (Irie et al., 2015; Magnúsdóttir et al., 2013). Overall, these findings underscore the evolutionary divergence of transcription factor regulatory networks in mouse and human PGCs, reflecting inherent differences in embryonic structure and developmental processes that likely arise from species-specific utilization of critical transcription factors.



Introduction Figure 1. The differences between the mouse and human germline during gastrulation highlight the origin of PGC induction. Diagrams of mouse and human gastrulating embryos depict the differences in the morphology of structure as well as the potential locations and origin/s of PGC specification. Furthermore, key differences in the transcription factor networks regulating this process are also highlighted. A↔P represents the anterior-posterior axis; (from Kobayashi and Surani, 2018).

2.2 Migration of human PGCs to the primary gonads

In the context of hPGCs, studies on E12–13 human embryos identified a cluster of these cells (Hertig et al., 1958), a finding recently supported by the detection of individual cells expressing key PGC marker genes in cultured human embryos (Chen et al., 2019). Furthermore, a single-cell transcriptomic study on a rare E16-17 human embryo reported a small number of hPGCs in the posterior region of the gastrulating epiblast (Tyser et al., 2020). Specified hPGCs are first observed in the early fourth week (approximately embryonic day - E24) in the extra-embryonic yolk sac wall near the allantois (**Introduction Fig. 2**), a location analogous to mPGCs at embryonic day 8 (Witschi, 1948). Subsequently, these cells migrate through the hindgut and dorsal mesentery and colonize the developing primary gonads (genital ridges) by the early sixth week (approximately embryonic day - E37, **Introduction Fig. 2**) (Witschi, 1948). During the migration phase, hPGCs undergo comprehensive epigenetic reprogramming, including global DNA demethylation and chromatin re-organisation (Tang et al., 2015, Ramakrishna et al., 2021). Once in the gonads during weeks 7-10, hPGCs, referred to as gonocytes (XY) or oogonia (XX), remain proliferative until around the tenth week, when they asynchronously enter either mitotic quiescence in males or meiotic prophase in females (Tang et al., 2016).



Introduction Figure 2. Cycle of human germline development, with the potential human primordial germ cell (hPGC) precursors highlighted in a black frame. After fertilization, the zygote develops into a blastocyst containing pluripotent pre-implantation epiblast cells, which give rise to all embryonic lineages, including the germline. Upon implantation, the blastocyst forms a bilaminar embryonic disc and undergoes gastrulation to generate the ectoderm, mesoderm, and endoderm germ layers. hPGCs are likely specified around embryonic day 17 (**highlighted in black frame**), although the exact timing remains uncertain. By week 4, hPGCs localize near the yolk sac wall, migrate through the hindgut, and reach the developing genital ridges. During migration, hPGCs undergo genome-wide epigenetic reprogramming, including DNA demethylation, to erase imprints and somatic epigenetic marks. In fetal development and adulthood, gonadal germ cells undergo meiosis and gametogenesis, accompanied by genome remethylation to establish epigenetic signatures essential for generating a totipotent zygote upon fertilization (modification of Tang et al., 2016).

3. Models of human primordial germ cells like cells (hPGCLCs) *In vitro*

Due to ethical and technical limitations that restrict direct studies on early human PGCs, researchers utilise *in vitro* models that are experimentally manageable for mechanistic investigations (Irie et al., 2015, 2018; Sasaki et al., 2015; Tang et al., 2016, 2022; Kobayashi et al., 2017). Although these models are derived *in vitro*, they can still be highly informative for germline biology when their transcriptomic profiles are comprehensively compared with rare human embryo samples and animal models, including *in vivo* development studies of nonhuman primates like marmosets.

In vitro studies in mice have been pivotal for understanding PGC development and gametogenesis, providing insights applicable to other models, including humans. *Ex vivo* studies of mouse embryos revealed that most pre-gastrulation epiblast cells (E5.5–E6.25) are germline competent (Ohinata et al., 2009), which led to the development of a two-step *in vitro* model for mPGC specification (Hayashi et al., 2011). In this system, naïve mouse embryonic stem cells (mESCs) are first differentiated into post-implantation epiblast-like cells (mEpiLCs). Secondly, when exposed to bone morphogenetic protein 4 (BMP4) and cytokines, these mEpiLCs efficiently generate mouse PGC-like cells (mPGCLCs) that closely resemble their *in vivo* counterparts (Tang et al., 2016). Remarkably, mPGCLCs can undergo gametogenesis, producing functional sperm and eggs after transplantation into gonads (Hayashi et al., 2011; 2012a).

Furthermore, the protocol for inducing mPGCLCs employs a combination of signalling molecules, including BMP4 and BMP8b, along with cytokines such as stem cell factor (SCF), leukemia inhibitory factor (LIF), and epidermal growth factor (EGF). These components have remained largely consistent across protocols for *in vitro* PGCLC induction from pluripotent stem cells (PSCs) in various species, including humans (Teague et al., 2023). This *in vitro* system has provided a robust platform for investigating the mechanisms underlying mPGC development and has facilitated the establishment of a comparable model for human PGC induction, which has only recently been developed.

The PGC-competent cells can be induced from cultured PSCs via two main approaches: (1) cultivating self-renewing cells under "4i" conditions with inhibitors for GSK3β, MEK, p38, and JNK (Gafni et al., 2013; Irie et al., 2015), and (2) generating transient pre-mesendoderm (PreME) populations (Sasaki et al., 2015; Kobayashi et al., 2017). In the first approach, Irie et al. (2015) demonstrated that human PSCs cultured under 4i conditions maintain high competence for hPGCLC fate by inducing the primitive streak-like markers Brachyury T and Heart and Neural Crest Derivatives Expressed 1 (HAND1), while also

exhibiting a presumptive posterior primitive streak feature (Mendjan et al., 2014). In contrast, when cells are cultured in conventional medium containing basic fibroblast growth factor (bFGF), this competence is progressively lost (Gafni et al., 2013; Gkountela et al., 2013). Moreover, while MAPK inhibitors can alter the epigenetic state of stem cells to potentially retain their pluripotency (Gafni et al., 2013), the molecular basis underlying hPGC competence remains to be fully elucidated (Irie et al., 2015). A similar competence is achieved from the second approach, where transient PreME induction is performed via a stepwise system rather than continuous culturing. In both approaches, when these competent cells, whether derived from 4i or PreME conditions, are aggregated into three-dimensional embryoid bodies, they can differentiate into PGCLCs (Irie et al., 2015; Sasaki et al., 2015; Kobayashi et al., 2017).

Despite relying on mouse *in vitro* differentiation protocols, comparative studies have revealed significant differences between hPSCs and their mouse counterparts with hPSCs occupying a unique pluripotent state distinct from that of mPSCs (Nakamura et al., 2016). This finding aligns with earlier research that positioned hPSCs transcriptionally between mESCs and mEpiLCs. As a result, hPSCs have been characterized as possessing an "extended primed pluripotency," a state that is distinct from both the naïve pluripotency of mESCs and the primed state of mEpiLCs. This extended primed state endows hPSCs with the competence to generate all cell types that emerge after implantation, including extraembryonic tissues like the amnion and germ cells (Saitou and Hayashi, 2021).

4. Interplay between WNT, BMP and Nodal signalling patways in mamalian PGC specification

In terms of signaling pathways mediating PGC competence and induction, studies of mouse and pig embryos, along with *in vitro* models inducing hPGCs from hPSCs, highlight that wingless-related integration site (WNT) and BMP signaling pathways are conserved and critical for PGC fate across mammalian species (Kobayashi et al., 2017; Ohinata et al., 2009). WNT signaling is essential for establishing the identity of the posterior epiblast and driving gastrulation and mesoderm formation, with WNT activators abundantly expressed in the posterior epiblast. Similarly, BMP signaling, particularly via BMP2 and BMP4, is crucial for inducing PGC specification (Kobayashi and Surani, 2018). In rodents, the extraembryonic ectoderm (**Introduction Fig. 1**) serves as a major source of BMP4, while the visceral endoderm is the site of BMP2 signalling. In contrast, non-rodent mammals such as pigs and rabbits rely on the posterior epiblast and mesoderm to express BMP4 via WNT at the early primitive streak stage, coinciding with the onset of gastrulation (Yoshida et al., 2016). However, primates, including humans, exhibit unique mechanisms. In cynomolgus monkey embryos, nascent PGCs have been shown to emerge from the early amnion, an extraembryonic membrane formed from the pluripotent epiblast shortly after implantation in primates (Sasaki et al., 2016), similarly as seen in humans (**Introduction Fig. 1**). Furthermore, in monkey embryos, WNT is expressed in the cytotrophoblast adjacent to the amnion, while the amnion itself produces BMP, potentially providing an additional environment that may support PGC specification (Sasaki et al., 2016). Therefore, despite structural and temporal differences in embryonic development between rodents, non-rodents, and primates, research underscores the shared roles of WNT and BMP signaling across species.

In terms of hPGC specification *in vitro*, Castillo-Venzor et al. (2023) explored the regulatory mechanisms of the transient state (PreME) that confers competence for PGC fate, utilizing temporal modeling, single-cell transcriptomics, and comparative analysis with primate animal models as well as rare human embryo transcriptomic atlases. Their analysis revealed that hPGC-competent cells exhibit a distinct signaling signature characterized by active Nodal (a key member of the TGF-β signaling pathway), and WNT pathways and low expression of BMP inhibitors like BAMBI and CER1, which likely facilitates PGC specification. Nodal plays a dual role in maintaining pluripotency in the epiblast and guiding its differentiation into various cell fates (Arnold and Robertson, 2009). Additionally, BMP4 expression is dependent on Nodal, as Nodal functions both downstream of BMP signaling and within a feedback loop to sustain BMP4 expression (Mulas et al., 2017). Furthermore, an increase in EOMES expression, a transcription factor as highlighted in **Introduction Fig. 1** crucial for hPGC competence (Chen et al., 2017; Kojima et al., 2017) was observed, while further activation of mesoderm factors hindered PGC specification (Castillo-Venzor et al., 2023). The precise interplay of signaling pathways, transcription factors, and intrinsic cellular heterogeneity aligns with the relatively small number (~100–200) of founder PGCs observed *in vivo* (Saitou et al., 2002; Kobayashi et al., 2017).

Depending on the trajectory of the *in vitro* differentiation protocol, hPGCLC-competent populations can be characterized by the expression of both amnion markers (e.g., TFAP2A induced by BMP signaling) and mesendoderm markers (e.g., TBXT induced by Nodal and WNT signaling) (Sasaki et al., 2016; Yokobayashi et al., 2017; Castillo-Venzor et al., 2023). al., 2023). In humans, BMP signalling (via BMP4) alone induces amnion formation, while WNT and Nodal induce the primitive streak; only the correct combination of these signals leads to PGC formation (Teague et al., 2023). Teague et al. (2023) proposed that these signaling pathways work together, guiding differentiating cells through a multidimensional Waddington epigenetic landscape. In this model, as described by Waddington (2015 edition), competence is not restricted to a linear trajectory tied to developmental stages, where cells can only move forward or backward. Instead, it is represented as a plateau, where cells can move in multiple directions, differentiating into various cell types, rather than acting separately in a purely permissive manner. In that respect, BMP directs cells toward the amniotic ectoderm, whereas WNT and Nodal steer them toward mesendoderm (ME) or primitive streak-like cells (PSLCs) (Teague et al., 2023). PGCs potentially reside between these states, and cells can reach the PGC fate via multiple pathways. For example, cells initially exposed to BMP move toward amniotic ectoderm-like cells (AELC), but timely exposure to Nodal and WNT can redirect them toward the PGC fate. Conversely, cells first exposed to WNT and Nodal move toward PSLCs, but subsequent BMP exposure before commitment can redirect them to become PGCs, similar to protocols involving an intermediate PreME step. From this perspective, there is no single "true" competent state or inductive signal; the appropriate signal depends on the cell's current state, a finding corroborated also by Alves-Lopes et al. (2023). This model reconciles apparent contradictions regarding the origin of PGCLCs, whether amnion-like or mesoderm-like, by suggesting they can originate from both (**Introduction Fig. 1**). This might explain *in vitro* findings and

the presence of PGCLCs in both amnion-like and primitive streak-like regions of human embryoids (Jo et al., 2022; Zheng et al., 2019), as well as their intermediate positioning in single-cell RNA sequencing data of BMP-treated hPSCs (Moon et al., 2019).

Advancements in understanding the nature of competent hPGC precursors have enabled researchers like Alves-Lopes et al. (2023) to derive hPGCLCs from resetting naïve pluripotent precursors. This approach not only improved differentiation efficiency but also promoted the progression of these cells toward migratory-like PGCs *in vivo*, supported by *in vitro* hindgut organoid co-cultures. These resetting precursors, which may resemble *in vivo* epiblast counterparts converging toward the hPGC fate, hold promise for more accurately recapitulating hPGC specification and development, leading to enhanced protocols for hPGCLC progression. Although the current working and applied hPGCLC specification protocols generate pre-migratory cells with different efficiencies and trajectories that transcriptionally resemble their *in vivo* pre-migratory counterparts, these models remain invaluable for studying human germline-associated diseases, including infertility.

5. Conserved Germline RBP: NANOS

As highlighted in **Introduction Fig. 1**, most studies have examined the roles of transcription factors and their signaling networks in PGC specification and *in vitro* differentiation across various animals and in humans. Overall, specification and maintenance of PGCs require proper transcriptome programming, likely involving the inactivation of genes and pathways that promote differentiation while selectively activating those necessary to confer and protect totipotent potential (see review, Lebedeva et al., 2018).

Although many genes are broadly transcribed throughout the embryo, mRNA and proteins are selectively expressed in germ cells, likely due to post-transcriptional regulatory processes. These include, but are not limited to, nuclear pre-mRNA splicing, polyadenylation, cytoplasmic mRNA localization, stability, translation, and degradation (Cooper et al., 2009). This regulation is mediated by RNA-binding proteins (RBPs) and regulatory small RNAs, which impact mRNA stability and protein expression (Shaw et al., 2010). Understanding the mechanisms of post-transcriptional regulation, especially via RBPs, and how they transfer information from parent to progeny at the mRNA level will deepen our knowledge of early development and elucidate how their misregulation can lead to infertility and disease (Albarqi and Ryder, 2023).

Nanos is a conserved post-transcriptional regulator RBP that was initially identified for its essential role in posterior pattern formation during embryonic development in *Drosophila melanogaster* (Lehmann and Nüsslein-Volhard, 1991). Beyond embryonic patterning, Nanos is crucial for embryonic germline development and maintenance of germline stem cells across a diverse range of metazoans (Jaruzelska et al., 2003; Tsuda et al., 2003; Baines, 2005; Lai and King, 2013). In mammals, from mice up to humans, the genome encodes three Nanos paralogs: **Nanos1, Nanos2,** and **Nanos3.** Homologs of Nanos are also present in several model organisms such as *Caenorhabditis elegans*, *Xenopus laevis*, and *Danio rerio* (for review, see De Keuckelaere et al., 2018). The function of Nanos paralogs in germ stem cells is conserved across species, from invertebrates to mammals. For example, Nanos2 and Nanos3 play critical roles in Mus musculus

(mouse) germline development (Tsuda et al., 2003), while NANOS3 is essential for germline development in Homo sapiens (human) (Julaton and Reijo Pera, 2011).

5.1 Structure and interactions of NANOS proteins

Structurally, Nanos proteins are characterized by a conserved C-terminal zinc finger domain (CCHC)2 and variable N-terminal and C-terminal segment/s that are predicted to be intrinsically disordered regions (IDR) due to their low sequence complexity accompanied by varying lengths (**Introduction Fig. 3**). The zinc finger (ZF) domain mediates RNA binding and interacts with proteins like Pumilio, a conserved partner that provides specificity for mRNA target recognition and binding (Arrizabalaga and Lehmann, 1999; Asaoka-Taguchi et al., 1999; Jaruzelska et al., 2003). A well-known function of Pumilio is the establishment of embryonic body patterning through the repression of maternal *hunchback* (*hb*) mRNA in the abdomen, a process that requires cooperation with Nanos (Nos) in *Drosophila* (Lehmann and Nusslein-Volhard, 1987, 1991; Murata and Wharton, 1995; Sonoda and Wharton, 1999; Wang and Lehmann, 1991). More recently, studies have shown that Nanos cooperates with Pumilio to enhance the RNA-binding specificity and affinity of Pumilio for its cognate motif in *hb* mRNA (Weidmann et al., 2016). Mechanistically, crystal structures of Nanos-Pumilio-RNA complexes revealed that Nanos acts as a molecular clamp, stabilizing the interaction between Pumilio and RNA (Weidmann et al., 2016). In this context, the C-terminal region of Pumilio undergoes conformational changes to form new contacts with both RNA and Nanos. Moreover, Nanos enhances Pumilio's repression activity and RNA-binding affinity by interacting with nucleotides upstream of the Pumilio recognition motif (PRE). This interaction allows Nanos to modify the specificity of the repression complex, enabling the repression of RNAs that Pumilio alone cannot stably bind. High-throughput sequencing and RNA-binding assays conducted by Weidmann et al. (2016) further demonstrated that Nanos diversifies Pumilio's RNA regulatory networks by expanding the range of target mRNAs.

Notably, additional nucleotides situated upstream of the PRE are also critical for *hb* mRNA regulation and are proposed to be recognized by Nanos (Edwards et al., 2001; Sonoda and Wharton, 1999). Although Arvola et al. (2017) identified that the ZF domain of Nanos can bind a Nanos Binding Site (NBS) in RNA (also termed NRE), they found that this binding occurs only when the downstream PRE is present. Crystal structures of Nanos's ZFs in complex with RNA and Pumilio in *Drosophila* further reveal how the tandem ZF domains form specific nucleotide-binding pockets (Weidmann et al., 2016). Nanos and Pumilio also work together to repress *cyclin B* (*CycB*) mRNA in PGCs and germline stem cells (Asaoka-Taguchi et al., 1999). Similar to *hb* mRNA, *CycB* mRNA contains NREs with PRE-like motifs.

In contrast, in mice, Nanos2 directly binds RNA with specificity to the consensus sequence AUKAAWU (with K = G or U, W = A or U), utilizing its two zinc-finger domain (Codino et al., 2021), with each zinc-finger recognizing three nucleotides (Choo and Klug, 1994). In contrast to *Drosophila* Nanos, which relies on the Pumilio protein to bind certain mRNAs such as *hb* (Sonoda and Wharton, 1999; Weidmann et al., 2016), murine Nanos2 independently recognizes and binds its RNA targets. In murine spermatogonial stem cells (SSCs), Nanos2 predominantly associates with these mRNAs through their 3′UTRs

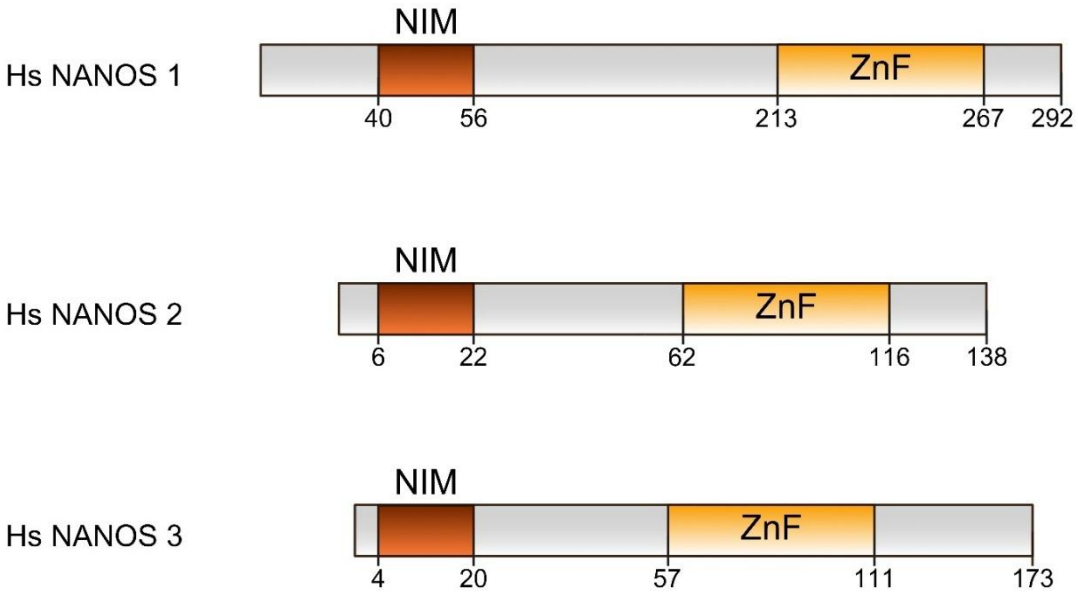
(Codino et al., 2021). The nature of mRNA targets of murine Nanos2 aligned with previous RNA immunoprecipitation (RIP) studies, which demonstrated that Nanos2 directly binds and represses specific mRNAs essential for SSC differentiation, such as Sohlh2, Dazl, and Taf7l (Zhou et al., 2015; Codini et al., 2021). Additionally, numerous other Nanos2 targets included mRNAs encoding proteins involved in cellular metabolism and biosynthetic processes, including ones involved in mechanistic target of rapamycin (mTOR) signaling and other pathways critical for SSC self-renewal regulation. Codini et al. (2021) highlighted that Nanos2 in mice selectively binds these key mRNAs, regulating their half-life and thereby controlling the metabolic and growth status of SSCs.

Overall, in *Drosophila*, the interaction between Nanos and Pumilio is stabilized by Pumilio interacting with its specific recognition motif, making it RNA-dependent (Sonoda and Wharton, 1999). However, in humans, NANOS1 interacts with PUMILIO2 independently of RNA (Jaruzelska et al., 2003). Similarly, mouse Nanos3 has been shown to interact with Pumilio in an RNA-independent manner (Lolicato et al., 2008). This indicates that while the Nanos-Pumilio interaction is conserved, the dependency on RNA for this interaction varies among species, and in certain species such as mice the mediated repression of Nanos2 is enacted directly on RNA targets.

While the partners of Nanos can vary depending on the organism and mRNA target, a common interacting partner is the CCR4–NOT deadenylase complex, which associates with the N-terminal region of Nanos proteins across diverse species (Kadyrova et al., 2007; Suzuki et al., 2010, 2012; Joly et al., 2013; Raisch et al., 2016). Although the N-terminal region is intrinsically disordered, the short 16-amino-acid CNOT1-interacting motif (NIM) embedded within the N-terminal of all three human NANOS paralogs (**Introduction Fig. 3**) is a structured and conserved domain. Furthermore, 17-amino-acid-long NIM peptides alone could recruit the CCR4–NOT complex to target mRNAs and initiate translational repression without causing mRNA degradation (Bhandari et al., 2014). The CCR4–NOT complex is composed of multiple independent modules assembled around the CNOT1 scaffold subunit (Collart and Panasenko, 2012; Wahle and Winkler, 2013), and it plays a pivotal role in post-transcriptional mRNA regulation by removing poly(A) tails, repressing translation, and promoting mRNA decay (Collart and Panasenko, 2012; Wahle and Winkler, 2013).

Murine Nanos2 has been previously shown to directly interact with the CNOT1 subunit of the CCR4–NOT complex. This interaction was mapped to be facilitated by a conserved 10-amino-acid region located at the N terminus of Nanos2 and the C terminus of CNOT1 (Suzuki et al., 2012 and Bhandari et al., 2014). In earlier work, Suzuki et al. (2010) used immunoprecipitation to highlight enrichment of components of the CCR4–NOT deadenylation complex that interact with Nanos2. They also demonstrated that the Nanos2/CCR4–NOT complex exhibits deadenylase activity *in vitro* and that specific meiosis-related RNAs, such as Sycp3, Stra8, Taf7l, Dazl, and Meisetz transcripts, were exclusively detected in NANOS2 protein precipitates and accumulated in its absence (Suzuki et al., 2010).

Notably, the three human and mouse Nanos paralogs (NANOS1–3) specifically interact with the C-terminal domain of CNOT1 (Suzuki et al., 2010 and Bhandari et al., 2014). Nanos1–3 bind to CNOT1 through the short NIM that is conserved across all vertebrates and some invertebrate species. Structural analyses of the human NANOS1 NIM peptide bound to CNOT1 reveal that the peptide inserts conserved aromatic residues into a hydrophobic pocket on the CNOT1 surface, effectively opening it. Substitution of these aromatic residues in the NIMs of NANOS1–3 disrupts their binding to CNOT1 and abolishes their ability to repress translation (Bhandari et al., 2014).



Introduction Figure 3. Domain organization of human (Hs) NANOS proteins (NANOS1–3). The domain organization of human NANOS1, NANOS2, and NANOS3 is depicted, showing their conserved C-terminal CCHC-type zinc finger domain (ZF, orange) and variable N-terminal and C-terminal extensions (gray). The NOT1 interaction motif (NIM, red), which is also conserved, is highlighted among the paralogs. Numbers below the protein diagrams indicate the residue positions marking the boundaries of the respective domain/motif. Figure adapted from Bhandari et al., 2014.

5.2 Functional significance of NANOS protein paralogues

In terms of functional significance, the NANOS family proteins execute distinct yet partially overlapping functions and exhibit specific expression patterns (Haraguchi et al., 2003; Jaruzelska et al., 2003; Kusz et al., 2009; Suzuki et al., 2007; Julaton and Reijo Pera, 2011). For instance, in a review by Lai and King (2013), the role of the Nanos/Pumilio complex across organisms highlighted repression of somatic gene expression, the cell cycle, and apoptosis, correlating perfectly with Nanos' function in germ cell development and survival.

In mice, Nanos1 is expressed in oocytes, the adult brain, and testes; Nanos3 is localized in PGCs; and Nanos2 is exclusive to male PGCs. Interestingly, mice lacking Nanos1 show no observable defects (Haraguchi et al., 2003), whereas the knockout of Nanos2 or Nanos3 leads to the loss of germ cells in males

or both sexes, respectively (Tsuda et al., 2003; Sada et al., 2009). Unlike Nanos2 and Nanos3, which are specifically expressed in germ cells (Tsuda et al., 2003), murine Nanos1 is predominantly found in the central nervous system (Haraguchi et al., 2003). Nanos1-deficient mice develop normally without noticeable differences from wild-type animals (Haraguchi et al., 2003).

6. Human NANOS1 association with infertility

In humans, the role of NANOS proteins in maintaining germ cell survival appears to be conserved, as all three paralogs are expressed in the gonads (Jaruzelska et al., 2003; Kusz et al., 2009; Julaton and Reijo Pera, 2011). Contrasting with the murine model, human NANOS1 is not expressed in the adult brain. NANOS1 mRNA and protein are expressed in embryonic stem cells and both fetal and adult gonads, including the adult ovary, which earlier studies had not reported (Jaruzelska et al., 2003; Julaton and Reijo Pera, 2011). Unlike mutations in NANOS2 (Kusz et al., 2009a) and NANOS3 (Kusz et al., 2009b), which have not been conclusively linked to male infertility, mutations in NANOS1 have been convincingly associated with this condition (Kusz-Zamelczyk et al., 2013; Janecki et al., 2020). Conversely, Wu et al. (2013) identified a novel mutation in the human NANOS3 gene linked to primary ovarian insufficiency in Chinese patients, revealing its critical role in destabilizing the NANOS3 protein and its relationship with PGC maintenance in female infertility.

Furthermore, critical insights have emerged from studies on hPGCLCs generated *in vitro* lacking PRDM14, a vital pluripotency and germ cell marker. These cells exhibited downregulation of several key genes related to hPGC function and pluripotency, including NANOS1 among other factors such as NANOG, LIN28A, and TRIM28 (Sybirna et al., 2020). This suggests that NANOS1 may play an important role in regulating germ cell biology in humans as early as the germ cell specification stage since the work by Sybirna et al. (2020) focuses on hPGCLCs that transcriptionally resemble pre-migratory *in vivo* counterparts. Mutations in the NANOS1 gene have been identified in infertile males with non-obstructive infertility but were absent in a control group of 400 fertile males, suggesting a significant role for NANOS1 in human male fertility (Kusz-Zamelczyk et al., 2013). Notably, the p.[Pro34Thr; Ser78del] mutation was associated with male infertility across two generations within a family pedigree. An exception was observed where the mutant allele was inherited from the father, indicating possible incomplete penetrance of this allele in that family. Importantly, female carriers of these mutations remained fertile.

The p.[Pro34Thr; Ser78del] mutation is located in the N-terminal region of NANOS1, which serves as an interface for interaction with GEMIN3 in germ cells (Ginter-Matuszewska et al., 2011). GEMIN3 is a component of the Survival Motor Neuron (SMN) complex, essential for the formation of small nuclear ribonucleoproteins (snRNPs) involved in accurate RNA splicing (Meister et al., 2000). It is also implicated in microRNA-mediated repression and is found in micro-ribonucleoprotein (miRNP) particles (Mourelatos et al., 2002). The interaction between NANOS1 and GEMIN3 occurs in the chromatoid body of germ cells, which is rich in microRNAs and regulatory components like Dicer and Argonaute proteins (Ginter-Matuszewska et al., 2011).

Functional assays using a quantitative yeast two-hybrid system demonstrated that the p.[Pro34Thr; Ser78del] mutation reduced interaction with GEMIN3 by 14%, whereas the p.[Pro34Thr] mutation alone, found equally in fertile and infertile men, did not affect this interaction (Kusz-Zamelczyk et al., 2013). Furthermore, the N-terminal region of human NANOS1 significantly differs from that of mouse Nanos1, including a unique 25-amino-acid sequence absent in mice. This sequence contains six serine residues, one of which is deleted in patients with the p.[Pro34Thr; Ser78del] mutation (Kusz-Zamelczyk et al., 2013). This divergence may reflect an evolutionary acquisition of specific reproductive functions by human NANOS1, such as the ability to bind proteins essential for germ cell development. The N-terminal regions of the paralogous proteins NANOS2 and NANOS3 are quite different from that of NANOS1, suggesting they cannot compensate for NANOS1 dysfunction caused by N-terminal mutations. This lack of redundancy could explain the dominant inheritance pattern of NANOS1 mutations. The association of NANOS1 mutations with Sertoli cell-only syndrome (SCOS), characterized by a complete absence of germ cells in the seminiferous tubules, implies that NANOS1 might play a crucial role in the self-renewal and maintenance of germline stem cells in men. To support this hypothesis, Janecki et al. (2020) investigated the role of human NANOS1 in germ cell apoptosis using the TCam-2 cell line, a human seminoma model representing germ cells prior to colonizing the gonads but also with aberrant transcriptome akin to cancer cells. Overexpression of wild-type NANOS1 in these cells reduced apoptosis by repressing the mRNA levels of several pro-apoptotic genes. In contrast, the infertility-associated double variant in NANOS1 p.[Pro34Thr; Ser78del] caused the protein to switch function, becoming pro-apoptotic and increasing cell death, as it impaired the repression of pro-apoptotic genes. Further supporting the findings of Janecki et al. (2020), Illaslan et al. (2022) reported that overexpression of NANOS1 induces a higher number of cells in the G1 and G2/M phases and a decrease of those in the S phase compared to control cells. This effect aligns with Gene Ontology (GO) analyses indicating that NANOS1 represses genes involved in the G1/S phase transition and in the p53-mediated DNA damage checkpoint during S phase. Downregulation of this checkpoint accelerates cell proliferation (Barr et al., 2017). These observations corroborate earlier results showing that NANOS1 downregulates GADD45 family members, which are key players in the G1/S phase DNA damage checkpoint (Janecki et al., 2020).

6.1 Human NANOS1 dysregulation in cancer

Germ cells and cancer cells share characteristics like self-renewal and rapid proliferation, with Nanos genes responsible for germline traits such as pluripotency and survival, traits also characteristic for tumour cells (De Keuckelaere et al., 2018). In *Drosophila* research model, aberrant expression of Nanos and other germ cell self-renewal genes can cause brain tumours, linking Nanos function directly to tumorigenesis (Janic et al., 2010). In human colorectal cancer cells, NANOS1 induces anti-apoptotic effects by downregulating pro-apoptotic kinases MAP3K1 and MAP2K3 (Miles et al., 2014). Additionally, NANOS1 overexpression is reported in several human breast cancer cell lines, causing downregulation of E-cadherin, a major epithelial cell-cell adhesion molecule and tumour suppressor (Strumane et al., 2006). Similarly, NANOS1

overexpression is associated with increased invasiveness in lung carcinoma and colorectal adenocarcinoma (Bonnomet et al., 2008).

Conversely, recent research by Rosemann et al. (2024) identified NANOS1 as a negative regulator of TGF-β signalling in oral squamous cancer cells. Functional studies showed that NANOS1 inhibits migration and invasion of oral cancer cells by interacting with TGFBR1 mRNA, promoting its decay, thereby inhibiting downstream TGF-β/SMAD signalling. Therefore, the effect of NANOS1 in cancer is context-dependent, exhibiting opposite roles in different cancer types.

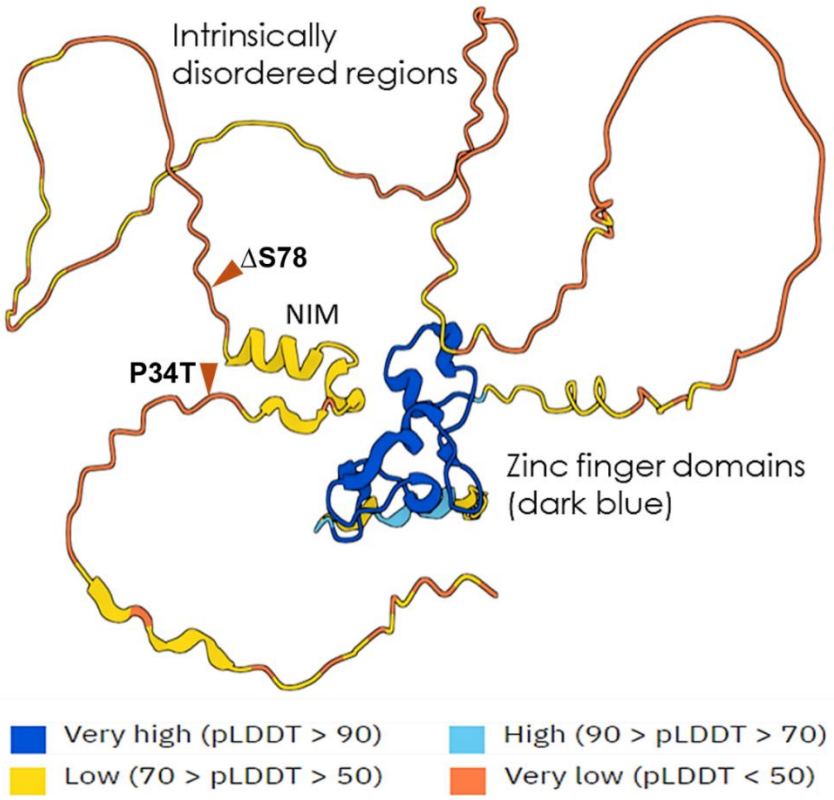
6.2 Assessing the potential impact of NANOS1 double variant on AlphaFold-predicted protein structure

According to AlphaFold predictions of the NANOS1 protein structure (**Introduction Fig. 4**), regions such as the zinc finger domain are predicted with very high confidence (predicted Local Distance Difference Test, pLDDT > 90). This high confidence indicates a well-structured domain critical for RNA binding and other molecular interactions, essential for the proper functioning of NANOS1. Similarly, the NIM is highlighted in yellow at a lower confidence level ($50 < pLDDT \leq 70$), reflecting lower structural confidence and implying higher flexibility. Despite this, the structure of the NIM region and its interaction with the CCR4–NOT deadenylation complex have been well-established through crystallography data (Bhandari et al., 2014).

In contrast, the variants P34T and Ser78del (ΔS78), which encompass the NIM domain (**Introduction Fig. 4**), are located in IDRs (pLDDT < 50) that lack stable tertiary structures. These disordered regions could play roles in recognition and dynamic interactions with other proteins or RNAs. However, predicting potential functional implications based solely on changes in structural flexibility is challenging. A loss of interactions of mutant NANOS1 can be hypothesized based on previously published data showing weakened interaction with GEMIN3, a protein partner (Kusz-Zamelczyk et al., 2013) or fewer RNA targets (Janecki et al., 2020). On the other hand, the fact that these mutations are situated in regions with significantly lower pLDDT suggests a higher degree of disorder, which could potentially translate to gain-of-function (GOF) or dominant-negative (DN) effects, as suggested by Gerasimavicius et al. (2022). The authors highlighted that DN and GOF mutations, which are less understood compared to loss-of-function (LOF) mutations, have distinct effects at the protein level.

GOF variants are often found in less ordered regions of proteins, where lower residue density results in fewer unfavorable energetic interactions. These mutations in disordered areas can promote abnormal interactions or inappropriate activation states, potentially increasing protein stability by stabilizing activated forms (Gerasimavicius et al., 2022). GOF mutations frequently exhibit stabilizing effects (21%) and may enhance a protein's ability to interact with other molecules or remain active when it should not. Additionally, dominant GOF mutations tend to cluster near functional domains, leading to aberrant activity without significantly disrupting the overall protein structure, in contrast to recessive LOF variants that are more uniformly distributed throughout the protein (Sivley et al., 2018). Given that the NANOS1 p.[Pro34Thr; Ser78del] variant exhibits a dominant inheritance pattern with penetrance limited to males, and that the double

mutation is located in a disordered region within a focal cluster near the functionally important NIM domain, it can be hypothesized that this variant may mediate a dominant GOF or DN effect on the protein structure. Such an effect could potentially induce unwanted activity or maintain the protein in an activated state during stages of germ cell development where its role might be minimal or not required. This aberrant activity might interfere with normal germ cell development and function, potentially leading to the observed infertility phenotype of Sertoli cell-only syndrome. Therefore, the location of the double variant and its inheritance pattern might support the hypothesis that a GOF or DN mechanism contributes to the disease pathology.



Introduction Figure 4. Structural prediction of the NANOS1 protein as generated by AlphaFold (last accessed September 2024). The model highlights intrinsically disordered regions and structured domains with predicted local distance difference test (pLDDT) scores. Zinc finger domain, critical for the protein's function, is shown in dark blue, indicating very high confidence (pLDDT > 90) in the structural prediction. Intrinsically disordered regions, shown in yellow and orange, represent areas with low (70 > pLDDT > 50) and very low (pLDDT < 50) confidence, respectively. The CNOT1-Interacting Motif (NIM) region (yellow, 70 > pLDDT > 50), although not predicted with high confidence, has a well-established structure based on published literature. Notable identified mutations P34T and ΔS78 are labelled for reference to the mutant NANOS1 counterpart. These mutations are located in the disordered parts of the protein, predicted with very low confidence (pLDDT < 50). The colour gradient corresponds to AlphaFold's confidence scale, as illustrated in the legend.

RESEARCH AIM AND OBJECTIVES

A double variant, p.[Pro34Thr; Ser78del], identified in infertile male patients, has been shown to impair the interaction of mutant NANOS1 with the DEAD-box helicase GEMIN3 (Kusz-Zamelczyk et al., 2013) and cause upregulation of some mRNAs encoding pro-apoptotic proteins (Janecki et al., 2020). Recent studies have demonstrated that this double variant influences proliferation, cell cycle progression, and apoptosis in the TCam-2 seminoma germ-cell-like model (Janecki et al., 2020; Illaslan et al., 2022), but its effect has not been fully elucidated in human PGCs. Therefore, the overarching goal is to achieve comprehensive insight into the structure and dynamics of the NANOS1 RNP interactome during the specification and early differentiation of hPGCs. We hypothesize that the double variant p.[Pro34Thr; Ser78del] modifies the NANOS1 RNP interactome at the stage of PGC specification. Therefore, this project seeks to compare the wild-type and mutant NANOS1 interactomes to identify potentially disrupted pathways critical for human germ cell specification and early development.

To achieve that, the following specific objectives were planned

1. Given the substantial differences in germ cell specification between humans and mice, the generation of an *in vitro* model of hPGC specification representing early development *in vivo* was of crucial importance. Therefore, a PGC differentiating model with temporally regulated expression of the wild-type and mutated NANOS1 protein respectively was planned.

2. A comprehensive global search for NANOS1-binding RNA targets in both the wild-type and mutant NANOS1 contexts. This type of broad, high-throughput analysis, employing enhanced crosslinking and immunoprecipitation (eCLIP), also provides the opportunity to identify the sites of NANOS1 interaction with RNA targets. (Studies of NANOS1-bound protein interactors for the wild-type versus mutant proteins were performed in parallel by another lab member).

3. RNA sequencing (RNA-seq) analysis of the transcriptome in both cell lines, wild-type and mutated, at different stages of hPGC specification to assess the altered RNA targets in a dynamic manner.

4. Functional studies of selected NANOS1-bound and altered RNA targets that differ between the wild-type and mutant cell lines.

Novelty

1. Post-transcriptional mechanisms of germ cell specification and development have not been extensively explored in mammals.

2. Furthermore, the field of post-transcriptional gene regulation involving the structure and dynamics of the RNP interactome in germ cell development remains largely unexplored, making this investigation particularly novel.

3. To date, the NANOS RNA interactome has not been explored in human PGCs but has been studied in model organisms and mice. Therefore, using an *in vitro* differentiation human PGC model that transcriptionally resembles *in vivo* PGCs is a novel approach for investigating post-transcriptional mechanisms and constructing the NANOS RNP interactome.

Model of study

A key aspect of this study lies in the experimental model used to generate hPGC-competent cells from cultured hESCs under defined *in vitro* conditions. This approach involves initially generating transient induced mesendoderm-like cells (iMELCs) or PreME precursors (Sasaki et al., 2015; Kobayashi et al., 2017). Subsequently, these precursors aggregate into embryoid bodies, from which they can differentiate into PGCLCs at efficiencies ranging from 10% to 40% in response to BMP and other cytokines (Irie et al., 2015; Sasaki et al., 2015; Kobayashi et al., 2017). The main cell line employed, the W15 (46, XY) hESC line, is endowed with a NANOS3-tdTomato fluorescent marker (Kobayashi et al., 2017) to track human PGCLC differentiation. This well-characterized differentiation model provides a crucial platform for examining the effects of genetic variants, such as the NANOS1 p.[Pro34Thr; Ser78del] double variant, which has been linked to male infertility. Importantly, the W15 hESC line is used to create dox-inducible NANOS1 cell lines, each expressing wild-type (wt) and mutant (mut) proteins respectively, and will be subjected to optimized differentiation protocols to successfully generate the respective PGCLCs. The latter will be sorted based on fluorescence markers and subjected to high-throughput technologies such as eCLIP and RNA-seq to identify the NANOS1 wt and mut bound and altered RNA targets.

Significance:

It is anticipated that the results of this project will advance the research field in several ways. First, the study will provide insights into the dynamics of post-transcriptional gene expression pathways regulated by NANOS1 during early human germ cell specification and development. Second, it will aim to establish a model of the NANOS1 RNP interactome in human PGCLCs during differentiation, enhancing our understanding of both healthy and diseased states. Third, identifying NANOS1-controlled RNP interactomes will lay the groundwork for future large-scale studies exploring RNP interactome complexity in germ cell specification and early development. Fourth, employing human germ cell models will improve the translational relevance of the findings, as key events differ between human and mouse germ cell specification. Finally, the elements of the NANOS1 interactome discovered through this research may be further investigated for their involvement in human infertility, potentially enhancing genetic diagnoses for infertile

couples undergoing *in vitro* fertilization and informing the development of novel therapeutic strategies. Notably, human germ cell differentiation models are critical experimental systems, as disruptions in these cells can have significant implications for progeny. By leveraging them, the project aims to generate findings that are more readily translatable to clinical contexts than those obtained from animal research models. The severe infertility phenotype in patients carrying the mutant NANOS1 (characterized by the absence of germ cells in seminiferous tubules) strongly suggests that the pathways influenced by NANOS1 may be crucial for human reproduction. In summary, this project seeks to deepen the understanding of NANOS1's role in early human germ cell development and its implications for reproduction and disease. Ultimately, these findings have the potential to enhance knowledge of human infertility and lead to more effective genetic diagnoses and therapies for affected individuals.

MATERIALS AND METHODS

Human embryonic stem cell culture Conventional undifferentiated hESCs the W15-NANOS3-T2A-tdTomato cell line (**Fig. S1**), 46 XY (Kobayashi et al., 2017), were cultured in Essential 8 (E8) medium with the associated supplement (Chen et al., 2011) on plates pre-coated with 1 µl/ml vitronectin for at least 1 hour. The media were replenished daily. Cells were either passaged in clumps using Versene (Lonza, cat. 17-711E) every 3-4 days or dissociated into single cells with Accutase (Thermo Scientific, cat. A1110501), counted, centrifuged, and seeded onto 6-well plates (Corning, cat. 3335). All reagents for E8 W15 culture were sourced from Thermo Fisher Scientific. For transgene induction in W15 cells or during differentiation, 500 ng/ml doxycycline (dox, Sigma, cat. D9891-1G) was added to the media as specified.

Pre-me and hPGCLC stage induction The induction of W15 into hPGCLCs was accomplished by differentiating W15 into mesendodermal precursors (pre-me) as outlined, with modifications, by Kobayashi et al. (2017). The protocol initiates with the preparation of aRB27 medium, which is composed of Advanced RPMI 1640 Medium supplemented with 1% (vol/vol) B-27 supplement, 0.1 mM non-essential amino acids (NEAA), 100 U/ml penicillin, 0.1 mg/ml streptomycin, and 2 mM L-glutamine (refer to **Tab. S1**). For mesendoderm (ME) induction, conventional hESCs are dissociated into single cells using Trypsin/EDTA (Lonza) and seeded onto vitronectin-coated wells at a density of 200,000 cells per well in a 12-well plate. The cells are then cultured in pre-me induction medium for 6 h for edited cell lines and for 12 h for the unedited control cell line. The ME induction medium comprises aRB27 medium supplemented with Activin A, 3 µM GSK3 inhibitor (CHIR-99021), and 10 µM ROCK inhibitor (refer to **Tab. S1**). Following preinduction, the cells are dissociated using Trypsin/EDTA (Lonza) and seeded into ultra-low attachment U-bottom 96-well plates (Corning, 7007) at a density of 4000–5000 cells per well in 100 µl of hPGC medium. The hPGC medium is supplemented with basal medium and cytokines, including BMP4, human LIF, SCF, and EGF, along with the ROCK inhibitor (refer to **Tab. S1** for detailed description). The aggregated cells form embryoid bodies (EBs) over a 4 – 5 day protocol during which the medium remains unchanged from the time of seeding, except for the addition of 500 ng/ml doxycycline (dox) (Sigma) to the medium on day 2 or day 3 of differentiation to induce protein overexpression in the respective cell lines.

Table S1. Composition of the culture media used during W15 differentiation toward hPGCLCs

aRB27 Basal Medium	Final concentration	Supplier
Advanced RPMI 1640	-	Thermo Fisher Scientific
B27	1%	Thermo Fisher Scientific
Nonessential Amino Acids	0.1 mM	Thermo Fisher Scientific
L-Glutamine	2mM	Thermo Fisher Scientific
Penicillin-streptomycin	100 U/ml (Penicillin) 0.1 mg/ml (Streptomycin)	Thermo Fisher Scientific
PreMe Medium, 6h preinduction	**Final concentration**	
aRB27 (as described above)	-	-
Activin A	100 ng/mL	Peprotech
CHIR-99021	3 µM	Sigma-Aldrich
ROCK inhibitor (Y-27632)	10 µM	Tocris
hPGC medium	**Final concentration**	
aRB27 (as described above)	-	-
BMP2	500 ng/ml	Peprotech
Human LIF	10 ng/ml	STEMCELL Tech.
SCF	100 ng/ml	Peprotech
EGF	50 ng/ml	R&D Systems
ROCKi inhibitor (Y-27632)	10 µM	Tocris

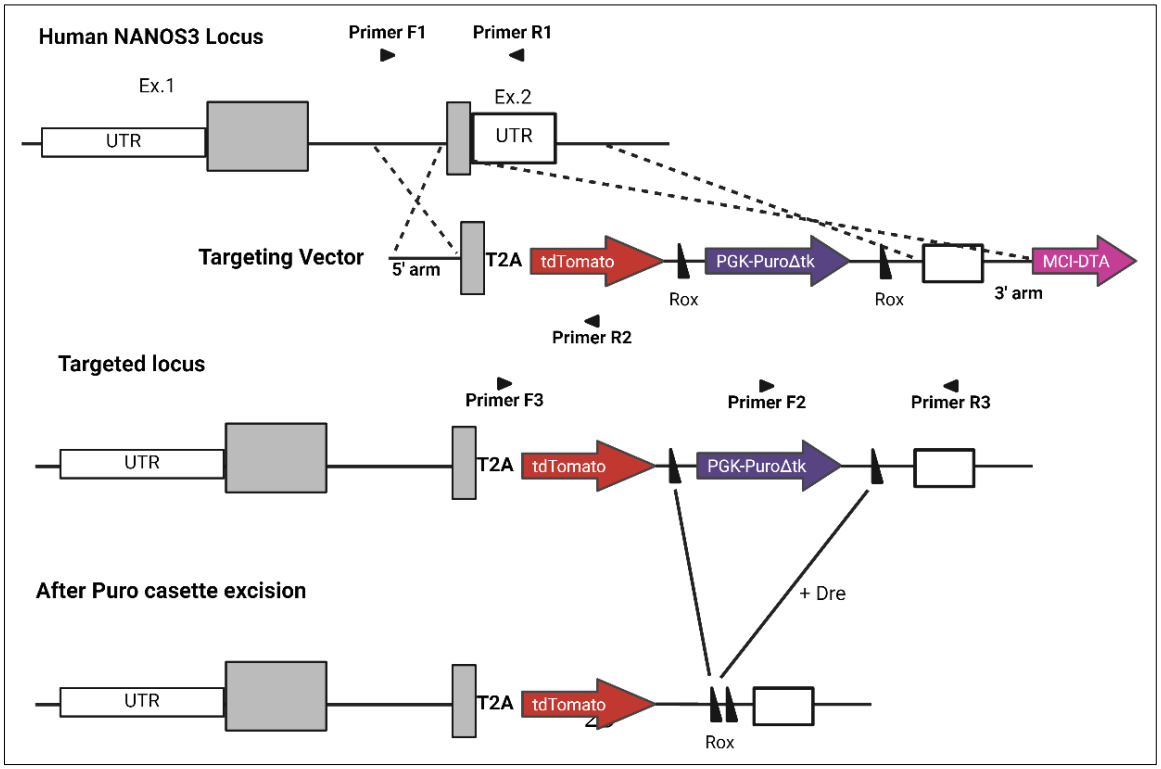


Figure S1. Targeting strategy for generating the NANOS3-tdTomato reporter W15 cell line, the primary model used in this project by (Kobayashi et al. 2017). CRISPR-Cas was employed to target the 3'UTR of exon 2 of NANOS3, incorporating the td-Tomato fluorescence marker sequence, the positive selection marker Diphtheria Toxin A (DTA), and the negative selection marker puromycin (PGK-PuroΔtk). The puromycin marker was subsequently removed through a second selection round by adding the Dre enzyme to the media, which excised the Rox-flanked sequences. A T2A (Porcine Teschovirus-1 2A) cleaving peptide mediates separate cassette translation. A detailed protocol for generating this edited cell line is provided by Kobayashi et al. (2017).

PiggyBac and Tet-On plasmid design and generation The PiggyBac system used to overexpress NANOS1 tagged with a 3x-FLAG tag and selection markers was designed to comprise two components: the transposon plasmid and the transposase (helper PBase). The transposon plasmid comprises the gene of interest flanked by inverted terminal repeats (ITRs), and a helper plasmid encoding the transposase enzyme, which recognizes these ITRs and facilitates the precise excision and insertion of the transposon into the host genome (Randolph et al., 2017) (**Fig. S2**). The key components of the transposon plasmid are shown in **S3**. For inducing NANOS1 overexpression in a non-continuous manner so as not to inhibit differentiation, a 3d generation tetracycline-responsive element promoter (TRE3G) was cloned into the PiggyBac transposon plasmid (**Fig. S3**, **Fig. S4A-B**). The PiggyBac-TET-ON system incorporates the PiggyBac plasmid into the genome at TTAA sequences (Park et al., 2018). This system is inducible due to the Tet-On 3G transactivator protein, which is constitutively expressed from $Promoter_{hPGK}$ and binds to $Promoter_{TRE3GS}$ only in the presence of dox. When dox is added to the culture medium, the transactivator protein undergoes a conformational change, binds to $Promoter_{TRE3GS}$, and activates the transcription of the downstream transgene (WT- and MUT-NANOS1-tag-reporter) (Zhou et al., 2006). In the first set of plasmids for both WT and MUT-NANOS1, the markers for selection post-transfection included mVenus fluorescence marker and puromycin antibiotic based selection (**Fig. S4A-B**). In the second set of plasmids, the markers of selection were simplified for a more efficient editing and maintenance of cell line pluripotency to solely fluorescence marker sfGFP. In **Fig. S5**, a diagram of WT- hNANOS1 showcases the changes. In both cases, 500 ng ml^{-1} dox is added to the media to induce hNANOS1 cassette overexpression. All the tested plasmids were synthesized by VectorBuilder and shipped in the format of *E. coli* glycerol stock. Plasmid DNA for each clone was then extracted and purified via the Qiagen Plasmid Maxi Kit (cat. 12163) per manufacturer's instructions to obtain a yield of 300 – 500 ug.

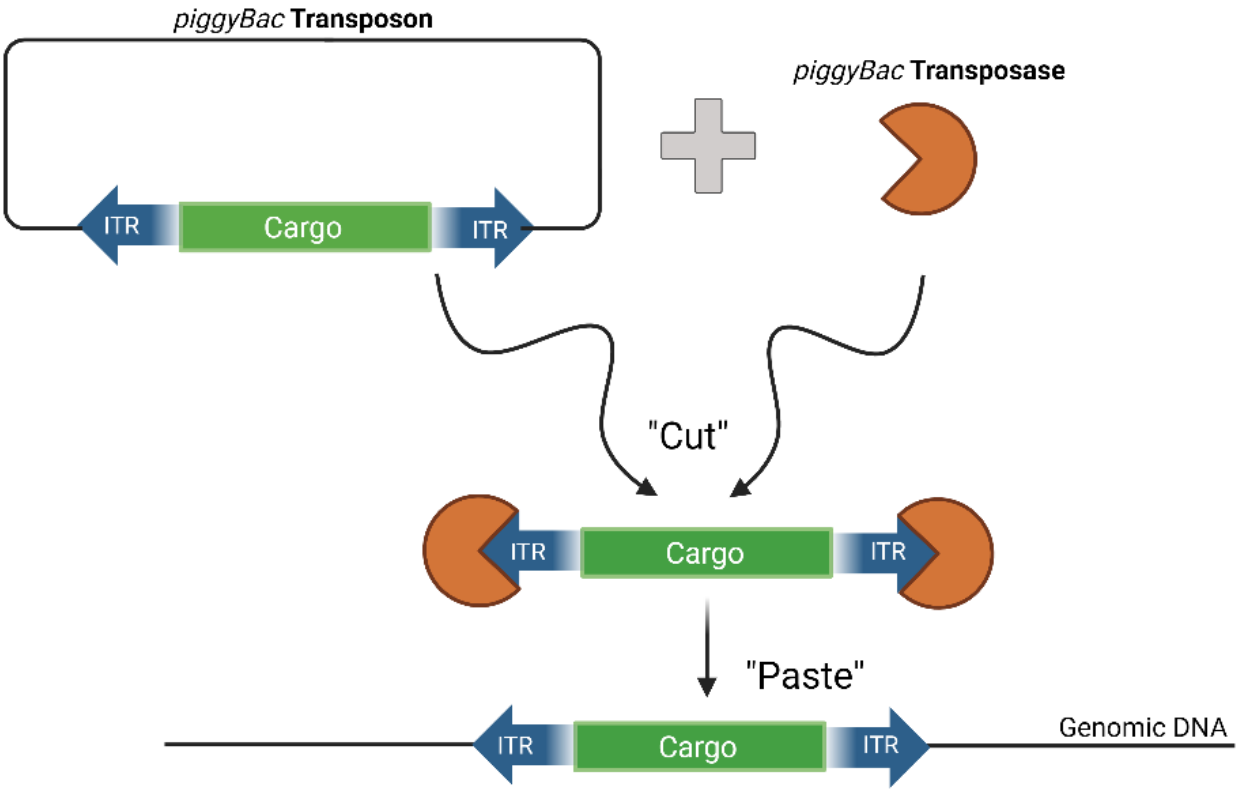


Figure S2. Diagram of the PiggyBac system featuring the transposon and transposase elements. The inverted repeats (ITRs) incorporated in the transposon allow for the transposase to cut the cassette cargo from the plasmid and paste it into the genomic DNA at TTAA sites. Adapted from https://blog.addgene.org/piggybac-ing-through-the-genome-editing-field (accessed December 14, 2024).

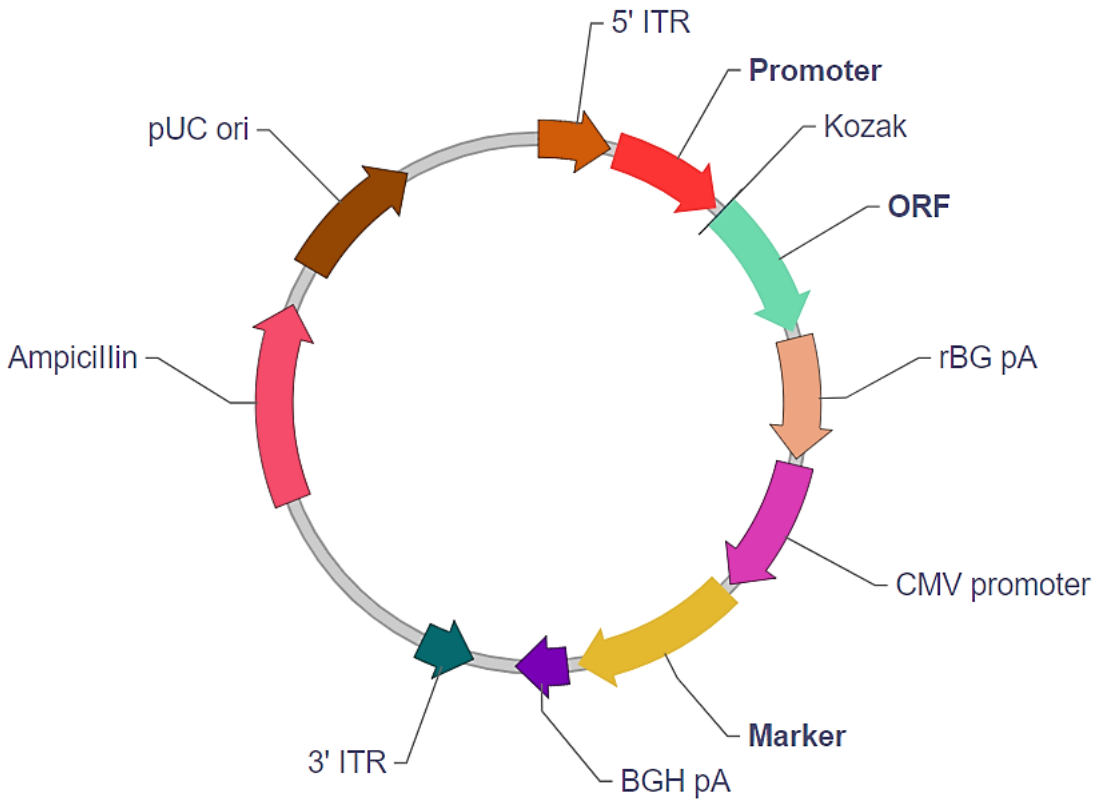
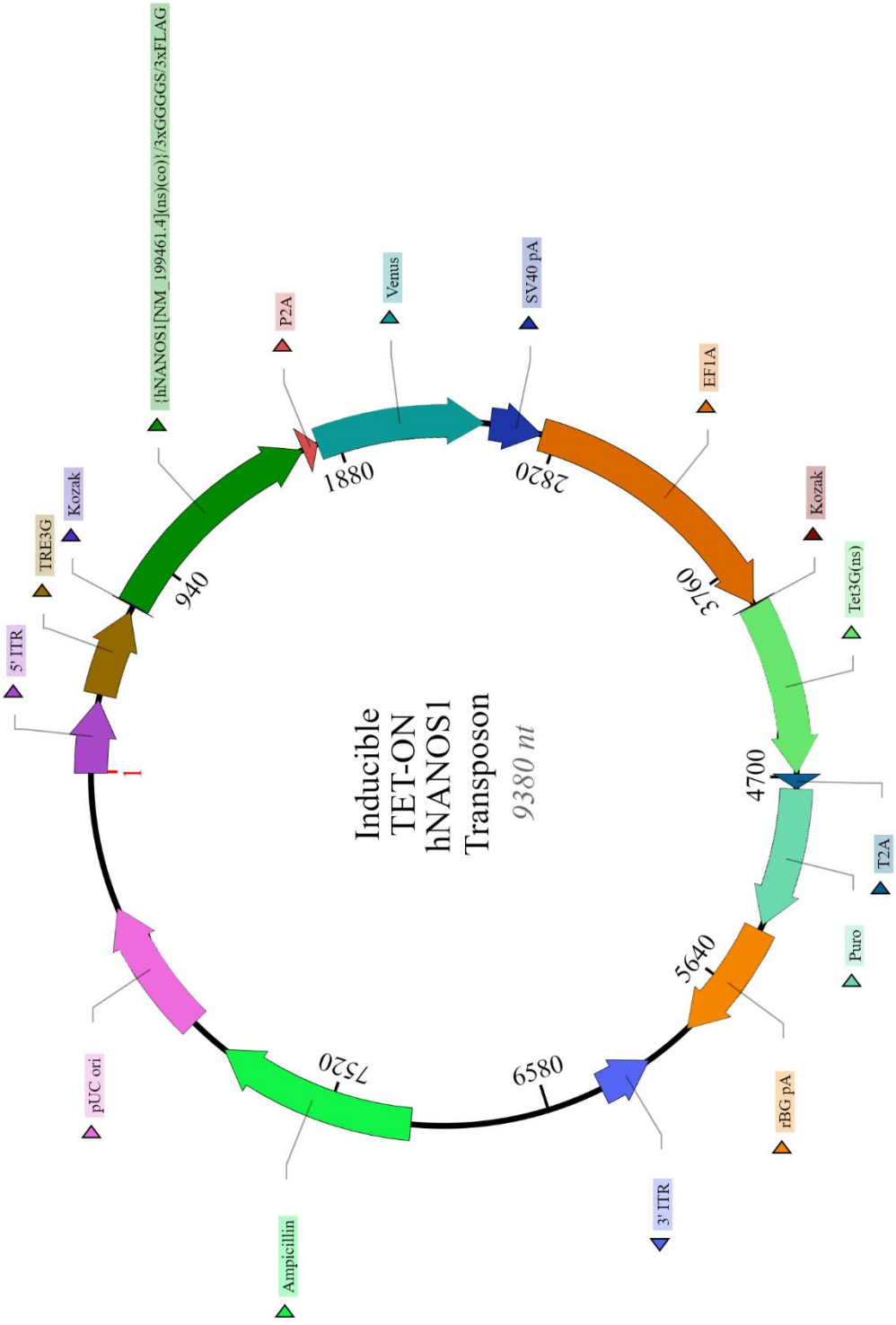


Figure S3. A classic transposon plasmid diagram from VectorBuilder (accessed 07.06.24, https://en.vectorbuilder.com). Key elements include: **5'ITR** – 5' inverted terminal repeat; **Promoter** of the gene of interest; **Kozak** sequence for translation initiation; **ORF** – open reading frame; **rBG pA** – rabbit ß-globin polyadenylation signal for transcription termination of upstream ORF; **CMV** promoter – human cytomegalovirus immediate early promoter; **Marker** - depending on applicability to allow selection and or visual tracing of cells; **BGH pA** – bovine growth hormone polyadenylation signal for transcription termination of upstream ORF; **3'ITR** – 3' inverted terminal repeat; ampicillin – resistance gene for plasmid maintenance via ampicillin selection in *E.coli*; **pUC ori** – pUC origin of replication for high copy number maintenance i



A

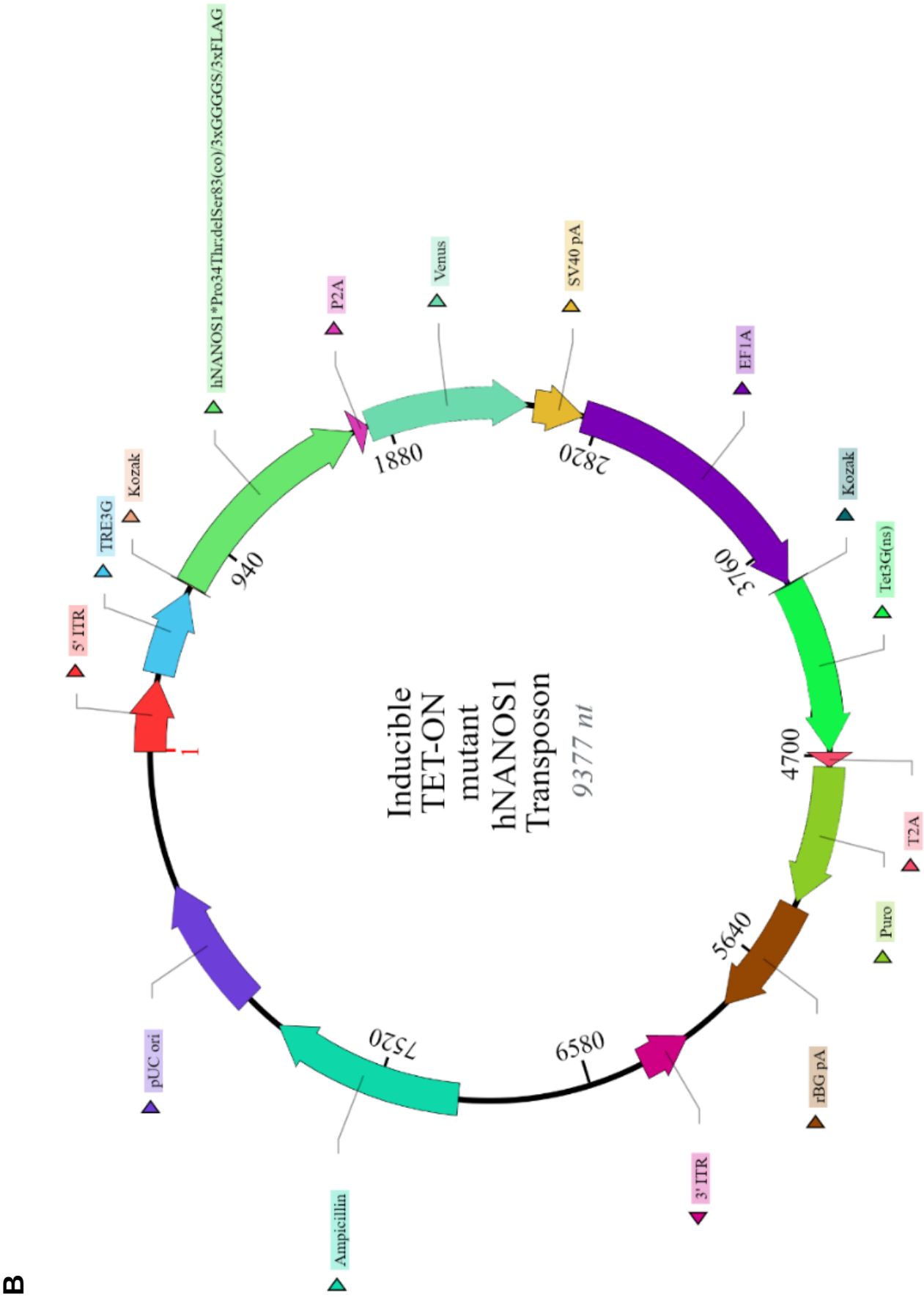

B
Inducible
TET-ON
mutant
hNANOS1
Transposon
9377 nt
1
940
1880
2820
3760
4700
5640
6580
7520
5' ITR
TRE3G
Kozak
hNANOS1*Pro34Thr;delSer83(co)/3xGGGGS/3xFLAG
P2A
Venus
SV40 pA
EF1A
Kozak
Tet3G(ns)
T2A
Puro
rBG pA
3' ITR
Ampicillin
pUC ori

Figure S4. Inducible hNANOS1 (human NANOS1) PiggyBac Tet-ON transposon diagram. A- The designed transposon incorporates elements of a PiggyBac transposon and a Tet-ON promoter (TRE3G), which is activated by the transactivator Tet3G in the presence of dox, resulting in the expression of the hNANOS1 sequence (hNANOS1[NM_199461.4]). This expression includes a 3xFLAG tag and an mVenus fluorescence marker, which is co-translated via the P2A (Porcine teschovirus-1 2A) cleavage peptide. The third-generation transactivator Tet3G sequence is located downstream of the EF1A promoter and is co-translated with puromycin through the T2A (Thosea asigna virus 2A) cleavage peptide for selection purposes. **B-** The inducible MUT-hNANOS1 transposon has the same features as the WT version, except it contains both variants (Pro34Thr; delSer83).

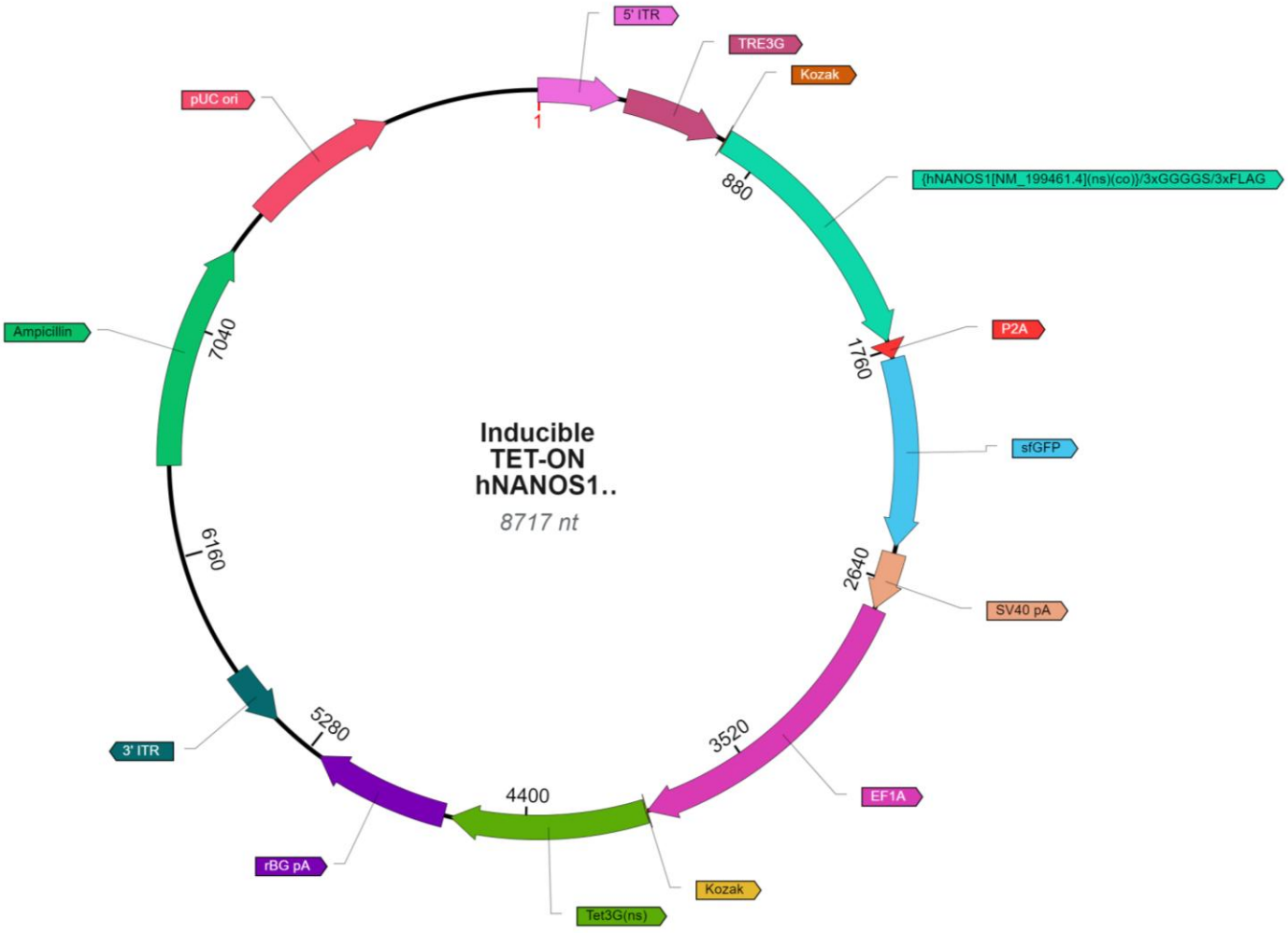


Figure S5. Inducible human NANOS1 (hNANOS1) PiggyBac Tet-ON transposon diagram with solely sfGFP marker for selection. The inducible hNANOS1 PiggyBac Tet-ON transposon diagram features a transposon with a Tet-ON promoter (TRE3G) activated by the transactivator Tet3G in the presence of dox. This leads to the expression of the hNANOS1 sequence (hNANOS1[NM_199461.4]), including the 3xFLAG tag and an sfGFP fluorescence marker, which is co-translated via the P2A (Porcine teschovirus-1 2A) cleavage peptide for sorting and selection. The Tet3G sequence is downstream of the EF1A promoter and is not accompanied by a selection marker.

Generation of CRISPR cell lines To detect endogenous NANOS1 (specific anti-NANOS1 antibody was not available), a 3x-FLAG tag (GACTACAAGGACCACGACGGCGATTATAAGGATCACGACATCGACT-ACAAAGACGACGATGACAAG) was inserted at the C-terminus of NANOS1 exon just before the STOP codon (TGA) using CRISPR-Cas9. The Alt-R™ HDR Design Tool (https://eu.idtdna.com/site/order/designtool/index/HDRDESIGN, accessed 03.07.24) was employed to design the CRISPR RNAs and the single-stranded oligo sequence for the HDR (Homology Directed Repair) donor. The guide and donor oligo sequences, along with their catalogue numbers and concentrations, are detailed in

Table S2. Each crRNA was initially tested for its specific ability to cleave the target site. Stem cells were cultured to 70 %-90 % confluency prior to electroporation. The CRISPR RNAs were resuspended with tracrRNA oligos (tagged with ATTO™ 488 nm fluorescence marker) in equimolar concentrations to achieve a final duplex concentration of 44 uM (gRNA). The duplex was heated at 95 °C for 5 min and then cooled at RT for 10 minutes. For electroporation, the Alt-R Cas9 enzyme was diluted to 36 uM in combination with Resuspension Buffer R (Neon System 10 ul Kit, Thermo Fisher Scientific, cat. MPK1025). For each well, the guide RNA and diluted Cas9 enzyme were mixed in equal volumes and incubated at RT for 20 minutes. The electroporation settings for the W15 hESC model were 1100 V + 30 ms (Pulse Width) delivered as 1 Pulse for 10 ul volume tips. During gRNA testing, the Alt-R Cas9 Electroporation Enhancer (IDT, cat. 1075915) was added to the final mixture at a working concentration of 10.8 uM. For the knock-in HDR, the enhancer was replaced by the single-stranded 3x-FLAG oligo (diluted to 100 uM stock, with a final working solution of 2 ul).

A total of 200,000 to 400,000 cells were required per 10 ul electroporation volume. Initially, cells were centrifuged at 150 x g for 10 min at RT. After removing the supernatant, the cells were washed in 5 mL of 1x PBS (Lonza, cat. 17-516F/12) and centrifuged again with the same settings. This second centrifugation was concurrent with the incubation of the CRISPR RNP. After supernatant removal, the cell pellet was resuspended in 9 ul Resuspension Buffer R per electroporation. Each electroporation involved adding 1ul CRISPR RNP complex, 9 ul cell suspension in R, and 2ul Alt-R Cas9 Electroporation Enhancer or single-stranded knock-in oligo to a 200 ul PCR tube. Following electroporation, the cells were transferred to pre-warmed wells in Essential 8 Media without antibiotics but with added ROCK Inhibitor to reduce cell death. 24 hrs before electroporation, when the cells reach approximately 70 % confluency, the E8 medium was supplemented with ROCK pathway inhibitor Y-27632 (final concentration 10 uM) to enhance cell survival post-electroporation. To increase editing efficiency after HDR, the medium was supplemented with nocodazole (Sigma Aldrich, cat. M1404-10MG) at a final concentration of 40 nM. For clonal generation, guide1 RNA (44 uM: crRNA_1 + trcRNA with ATTO 448nm) combined with the respective HDR oligo (11 uM) was used in 6-well plates, treated both with and without nocodazole.

Table S2. List of CRISPR 3x-FLAG knock-in reagents. These components were designed and synthesized by IDT (https://www.idtdna.com), including sequences and concentrations of CRISPR RNAs targeting the NANOS1 3' end prior to the STOP codon and the positive control Hs AAVS1, along with their HDR donor oligos. The positive control gRNA targets the human (Hs) AAVS1 safe harbour locus to create a CRISPR-Cas9–mediated double-stranded break. In stem cells that have undergone HDR, the provided donor template of positive control kit introduces a six-base insertion (indicated in red as an EcoRI recognition sequence) at the cleavage site for further restriction enzyme-based PCR assay for validation.

Design crRNA	Sequence	PAM	Strand	On-Target Score	Off-Target Score
Alt-R® CRISPRCas9 crRNA, 2 nmol_1 (IDT)	CTGGCAAGAAGCTGCGCTGA	AGG	+	53	44
Alt-R® CRISPRCas9 crRNA, 2 nmol_2 (IDT)	AGAAGCTGCGCTGAAGGCCC	GGG	+	36	21

Design HDR Donor Oligo	HDR donor sequence (+)	HDR donor sequence (-)	Enzyme Type	Transcript	Tag name(s)
Alt-R™ HDR Donor Oligo, 2 nmol_1 (IDT)	CCGCAGCGCCAGGGACGGCCCGCCTGGCAAGAAGCTGCGCGACTACAAGGACCACGACGGCGATTATAAGGATCACGACATCGACTACAAAGACGACGATGACAAGTGAAGGCCCGGGCTCCCGGCCGCCCAGGGTCGCCGCCGCC	GGCGGCGGCGACCCTGGGCGGCCGGGAGCCCGGGCCTTCACTTGTCATCGTCGTCTTTGTAGTCGATGTCGTGATCCTTATAATCGCCGTCGTGGTCCTTGTAGTCGCGCAGCTTCTTGCCAGGCGGGCCGTCCCTGGCGCTGCGG	S.p. Cas9	NM_199461.3	3xFLAG (+) strand
Alt-R™ HDR Donor Oligo, 2 nmol_2 (IDT)	CCGCAGCGCCAGGGACGGCCCGCCTGGCAAGAAGCTGCGCGACTACAAGGACCACGACGGCGATTATAAGGATCACGACATCGACTACAAAGACGACGATGA	GCGAGGGGCGGCGGCGACCCTGGGCGGCCGGGAGCCCGGGGCCTTCACTTGTCATCG	S.p. Cas9	NM_199461.3	3xFLAG (+) strand

	CAAGTGAAGGCCCGGGC TCCCGGCCGCCCAGGGT CGCCGCCGCCCCTCGC	TCGTCTTTGT AGTCGATGTC GTGATCCTTA TAATCGCCGT CGTGGTCCTT GTAGTCGCG CAGCTTCTTG CCAGGCGGG CCGTCCCTGG CGCTGCGG			

Design Hs AAVS1	Sequence	Primer Forward	Primer Reverse	Fragment 1	Fragment 2
Alt-R® CRISPRCas9 crRNA, 2 nmol	CCTCTAAGGTTTGCTTACGA	GCCAAGGACTCAAACCCAGA	CCCCGTTCTCCTGTGGATTC	561 bp	476 bp
Alt-R™ HDR Donor Oligo, 2 nmol	AGCCATCTCTCTCCTTGCCAGAACCTCTAAGGTTTGCTTA*GAATTC*CGATGGAGCCAGAGAGGATCCTGGGAGGGAGAGCTTGGCA	GCCAAGGACTCAAACCCAGA	CCCCGTTCTCCTGTGGATTC	561 bp	476 bp

*GAATC = EcoRI recognition site

DNA extraction, amplification and T7 Assay To screen for effective CRISPR guide RNA cuts, cells seeded in 24-well plates were collected 24-48 h post-transfection. Lucigen DNA Quick Extract Lysis Buffer (Biosearch Technologies, cat. QE09050) was added directly to cell pellets (50 µl for 96-well, 100 µl for 48-well, or 200 µl for 24-well plates). The following PCR program was run 65°C for 15 min, 68°C for 15 min, 98°C for 10 min, and then held at 4°C for storage. For genotypic PCR, 500 ng of extracted genomic DNA was used in a 50 µl reaction. Primers and PCR conditions for the nested PCR prior to the T7 assay are detailed in **Tab. S3**. Genotyping PCRs, utilizing the Advantage® 2 Polymerase Mix (Takara Bio, cat. 639201) and Herculase II Fusion DNA Polymerase (Agilent Technologies, cat. 600677), are also shown in **Tab. S3**. The T7 assay kit (New England Biolabs, cat. M0302S) was used to assess the cutting efficiency of each guide RNA, following the kit's protocol: PCR products were annealed and treated with T7 endonuclease I at 37°C for 15 min, generating fragments as depicted in **Fig. S6**. Resulting clones were genotyped for the presence of FLAG via PCR and confirmed by Sanger sequencing.

Table S3. Validation conditions required to genotype W15 hESCs following CRISPR Knockout. Primer sequences, PCR conditions, and product lengths for genotyping stem cells following guide RNA cut efficiency testing (PCR1 + PCR2, nested PCR), including positive control [Pos (+) Cont.] human AAVS1 and FLAG knock-in clone genotyping.

PCR	Forward primer	Reverse primer	Product length	GC-Melt (Takara Bio Kit)	DMSO % (Herculase II Fusion DNA Polymerase)	Annealing Temperature
PCR1	GGCGCTCTACACCAC CCATA	GGGGAATACCCGCC CAAATTA	821 bp	0.5 M	-	61°C
PCR2 (Nested)	AACGCGCACACCATC AAGTA	TAAGGGTGATGCTA CCAGCG	667 bp	1.5 M	-	62.9°C
AAVS1 Pos (+) Cont.	GCCAAGGACTCAAA CCCAGA	CCCCGTTCTCCTGTG GATTC	1037 bp	-	4%	57.5°C
Knock-In FLAG	GCACACCATCAAGTA CTGCC	GACGTCCCATGTCGA GCC	223 bp	0.5 M	-	63.1°C

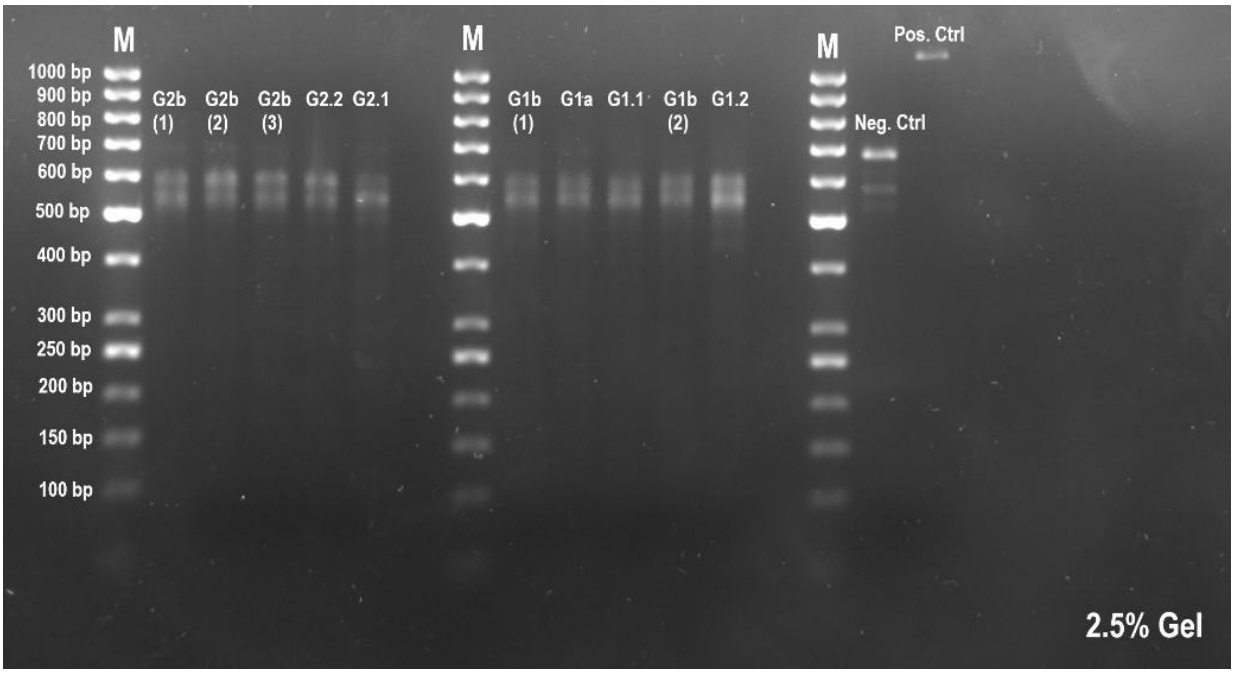


Figure S6. Genotyping of nested PCR products following the T7 assay. Guide 1 and Guide 2 were tested in two technical replicates without nocodazole (1.1, 1.2, 2.1, 2.2) and with nocodazole (conditions a and b, representing different biological replicates, with technical replicates indicated underneath by numbers). Only products with sufficient DNA were run on the gel. The negative control (667 bp) and positive control (1037 bp) without nocodazole are also shown. The second product for all guides is not visible due to diffusion. An unspecific band around 600 bp, likely from nested PCR and T7 assay, is visible in all samples including the negative control. Guide RNA 1 is expected to generate two products (105 bp and 562 bp) post-cut, as is guide RNA 2 (111 bp and 556 bp). The higher band is visible for both guides, with slightly stronger intensity for guide 1, especially in the G1.2 replicate.

Fluorescence activated cell sorting (FACS) PiggyBac clones For clones containing the fluorescence marker sf-GFP (see **Fig. S5**), the sorting protocol followed the method outlined by Park et al. (2018). Various ratios of transposon to transposase were tested, as detailed in **Tab. S4**, to determine optimal conditions for the W15 cell line, gene expression levels, transgene leakage, and differentiation efficiency. For the FACS protocol, 1 x 10^6 cells were plated in two 6-well plates. This high density was necessary to improve viability and reduce the tendency for W15 cells to self-differentiate. After electroporation of the transposon and transposase plasmids, cells were plated in E8 media with ROCK Inhibitor. The sf-EGFP marker, used for sorting into 96-well plates, was induced with dox 24 hrs post-electroporation. Cells were allowed to grow for 4-5 days to eliminate non-integrated products and recover from electroporation stress. Two days post-transfection, cells from both 6-well plates were combined into a 10 cm tissue-culture dish (Corning FALCON®, cat. 353003) to prevent overgrowth, with continued dox induction. Prior to sorting, 96-well flat bottom plates (Corning FALCON®, cat. 351172) were coated with Vitronectin (Life Technologies, cat. A14700) diluted in PBS to a final concentration of 0.5 μg/cm² and incubated for 1 h at 37ºC. Vitronectin was then aspirated and replaced with 200 μl E8 medium containing 10 μM ROCK Inhibitor, and plates were stored at 37ºC. Transfected cells were washed with 1x PBS, treated with Accutase for 2-5 min at 37ºC until detached, resuspended in E8 medium, transferred to 15 ml conical tubes, and centrifuged for 5 min at 400 x g at RT. After removing the supernatant and resuspending in media with ROCK Inhibitor, cell preparations were passed through a 40 μm cell strainer into FACS tubes to remove clumps and debris. Following filtering, cells were resuspended in FACS buffer without sodium azide: 1x PBS, 5% FBS (Hyclone, cat. 30160.03), and 2 mM EDTA (Merck Millipore, cat. 324506). Cells were kept on ice for no more than 2 h post-resuspension. An 85 - 100 μm nozzle was used during sorting to reduce sheath pressure, and a sorting speed under 2000 events per second on a BD FACSAria™ III Cell Sorter was employed to ensure accurate single-cell deposition and minimize cell death. While single-cell sorting was prioritized for precise editing, some wells were seeded with 3 cells to enhance the viability and stemness of the resulting cell lines (**Fig. S7**). Cells were first gated by FSC-A vs SSC-A and then by FSC-A vs FSC-H and SSC-A vs SSC-H to eliminate doublets (**Fig. S8**). Only cells expressing the dox-induced marker (sf-EGFP) were selected (**Fig. S7** and **Fig. S8**). After sorting, plates were incubated at 37ºC. Media was not replaced until 3-5 cell colonies were visible in the wells (approximately 3 days) to prevent colony loss. Media was then changed every 2 days until colonies were ready for passaging (approximately 8-10 days). Clones were passaged into larger wells up to 6-well plates before being stored in liquid nitrogen for further characterization and downstream analysis.

CRISPR FLAG knock-in cell lines The utilization of Alt-R™ CRISPR-Cas9 tracrRNA, ATTO™ 488 (IDT, cat. 10007810), and custom-designed crRNAs to assemble the ribonucleoprotein (RNP) for targeting 3x-FLAG Knock-in necessitated the application of FACS in a time-dependent manner (**Fig. S9**). To enhance outcomes, transfected cells underwent a wash with PBS containing 1% FBS before FACS, aimed at reducing non-specific binding and increasing the yield of positive cells. A negative control, involving the combination

of crRNA with a tracrRNA lacking fluorescence, facilitated the application of specific gating. Furthermore, to optimize results, FACS was conducted 24 hours post-transfection, ensuring the visibility of fluorescence at the appropriate time. Subsequently, the process of plate preparation, involving single cell seeding onto Vitronectin-coated plates, mirrored the methodology employed for generating PiggyBac Clones on the BD FACSAria™ III Cell Sorter (**Fig. S8** and **Fig. S9**).

FACS for bulk RNA sequencing On days 4-5, EBs were harvested from 96-well U-bottom culture plates and collected in a 15 mL Falcon tube. After a brief centrifugation at 400 x g for 1 min, the EBs were rinsed with 1x PBS and then subjected to a 3.5-min centrifugation at 400 x g. Subsequently, a quantity of freshly thawed Accutase ranging from 7.5 to 10 µl was added per EB. The EBs in Accutase were enzymatically digested at 37°C in a Heating Shaker (Lab Companion, CBS-350) at 850 rpm. Vigorous pipetting of the samples every 7 min expedited the digestion process, with this step repeated for a maximum of 3-4 rounds to prevent excessive digestion of the EBs. Following digestion, inactivation of Accutase was achieved by adding a 3x volume of basal media aRB27 (**see Tab. S1**). Cell counting was performed using the Countess™ 3 FL Automated Cell Counter (ThermoFisher Scientific, AMQAF2000) to adjust the sorting buffer volume accordingly. After centrifugation at 400 x g for 5 min, the supernatant was aspirated, and the cells were resuspended in cold sorting buffer and filtered before sorting into 1.5 ml Low Binding Microcentrifuge Tubes (ThermoFisher Scientific, cat. 90411). The sorting gates and mode are detailed in **Fig. S10** and **Fig. S11**.

The inclusion of the tdTomato fluorescence marker, previously introduced to the W15 unedited cell line via CRISPR at the 3' end of NANOS3, facilitated the sorting of uninduced Day 4 digested cells into positive and negative germ cell populations (Kobayashi et al., 2017). Sorting of the cells was carried out using the S3e™ Cell Sorter (488/561 nm – 100 mW lasers) (Bio-Rad, cat. 12007052), equipped with 4 fluorescence detectors with filters and the AutoGimbal™ System. The FL2 channel and 561 nm laser were employed to separate the cells into tdTomato-positive and -negative populations in non-induced controls for each analyzed cell line (WT-NANOS1 1.2.51 and MUT-NANOS1 3.2.5.2, as well as the unedited 15 cell line) (refer to **Fig. S10**).

In instances where dox was administered (500 ng/mL) on Day 3 of differentiation and continued until Days 4-5, automatic compensation was implemented using the ProSort™ Software, and two lasers were combined: 488 nm for sf-EGFP-NANOS1 overexpression and 561 nm for the tdTomato germ cell marker (channels FL1 and FL2). Similar configurations were applied when sorting cells using the Sonny MA900 Multi-Application Sorter and Software. Positive cells were double-gated, as illustrated in **Fig. S11**, while negative soma cells were single-gated for sf-EGFP-NANOS1. The FITC channel was utilized to visualize the sf-EGFP marker, and the PE channel was used to visualize the td-Tomato endogenously expressed germ cell marker (**Fig. S11**). Subsequently, the samples were centrifuged at 10,000 rpm for 1 min, the supernatant was removed, and the pellets were resuspended in Trizol (ThermoFisher Scientific, cat. 15596026) and stored in dry ice for further extraction and sequencing. The previously mentioned markers were regularly examined

during the clonal validation process and to assess differentiation efficiency using flow cytometry with the CytoFLEX Flow Cytometer (Beckman Coulter, catalog number B53000).

Table S4. Plasmid concentrations and the corresponding transposase-to-transposon ratios for both WT- and MUT-NANOS1. A total of 1 x 10^6 cells were electroporated per transfection using 100 µL tips from the Neon™ Transfection System 100 µl Kit (cat. MPK10025) in 120 µl of R Buffer. Electroporation was conducted under the following conditions: 1100 V, 30 ms, 1 pulse.

	hyPBase [µl]	Transposon NANOS1 WT [µl]	Transposon NANOS1 MUT [µl]
Concentration [µg/mL]	0.635	3.05	2.97
1:1 (2.4 µg : 2.6 µg)	4.54	1.02	1.05
1:2.5 (0.5 µg : 1.25 µg)	0.94	0.49	0.51
1:5 (0.5 µg : 2.5 µg)	0.94	0.98	1.01
1:10 (0.5 µg : 5 µg)	0.94	1.97	2.02

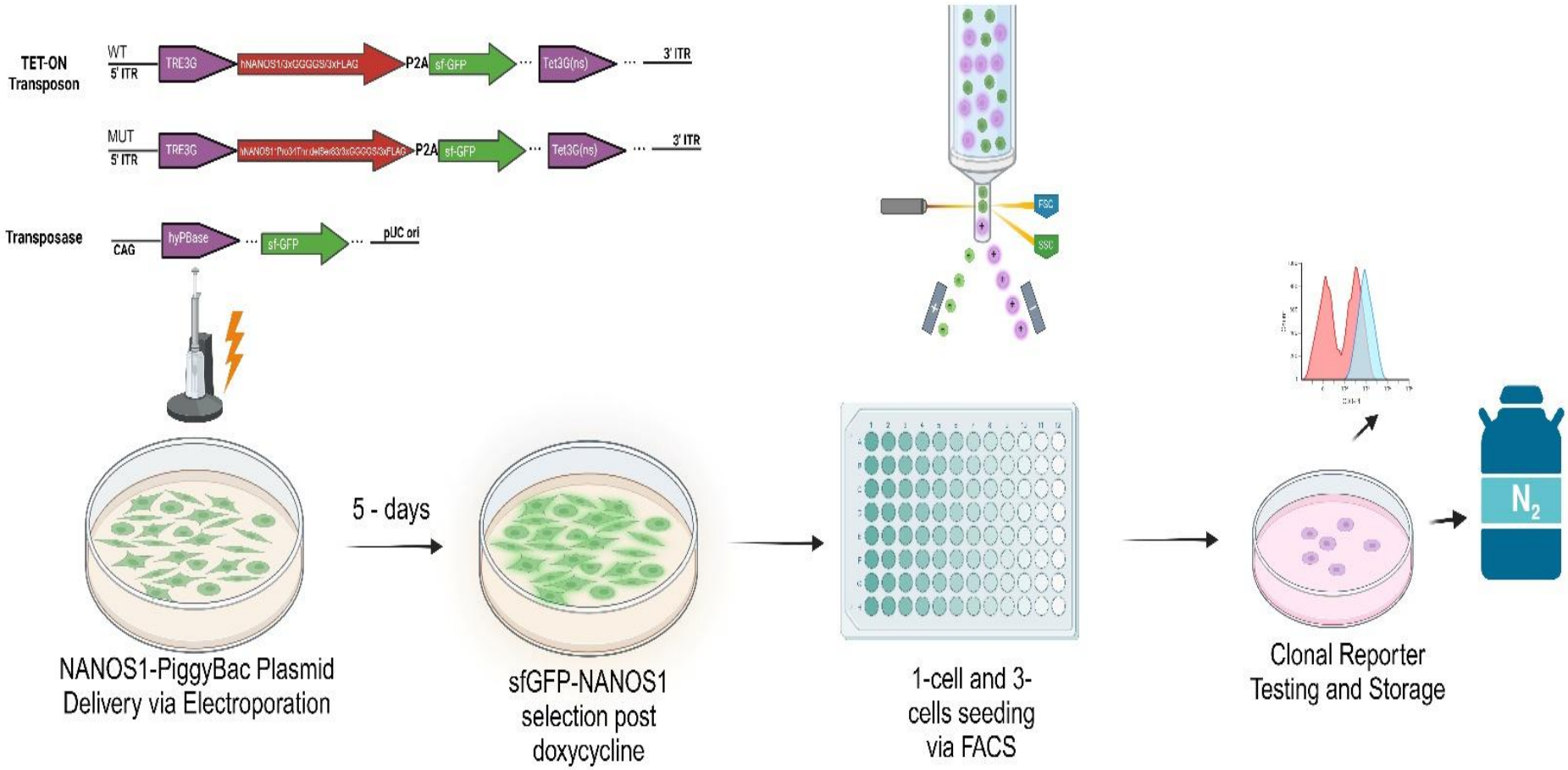

TET-ON Transposon
WT
5' ITR
TRE3G
hNANOS1(3xGGGGS)3xFLAG
P2A
sf-GFP
Tet3G(ns)
3' ITR
MUT
5' ITR
TRE3G
P2A
sf-GFP
Tet3G(ns)
3' ITR
Transposase
CAG
hyPBase
sf-GFP
pUC ori
FSC
SSC
5 - days
NANOS1-PiggyBac Plasmid Delivery via Electroporation
sfGFP-NANOS1 selection post doxycycline
1-cell and 3-cells seeding via FACS
Clonal Reporter Testing and Storage
N2

Figure S7. Schematic representation of cell line generation using PiggyBac and Tet-ON transposon for both WT and MUT cassettes including the hyPBase transposase introduced via electroporation. The process entails induction of sf-EGFP with doxycycline, FACS sorting by 1-cell or 3-cell seeding, cell line expansion followed by storage in liquid nitrogen and further analysis of signal expression via flow cytometry.

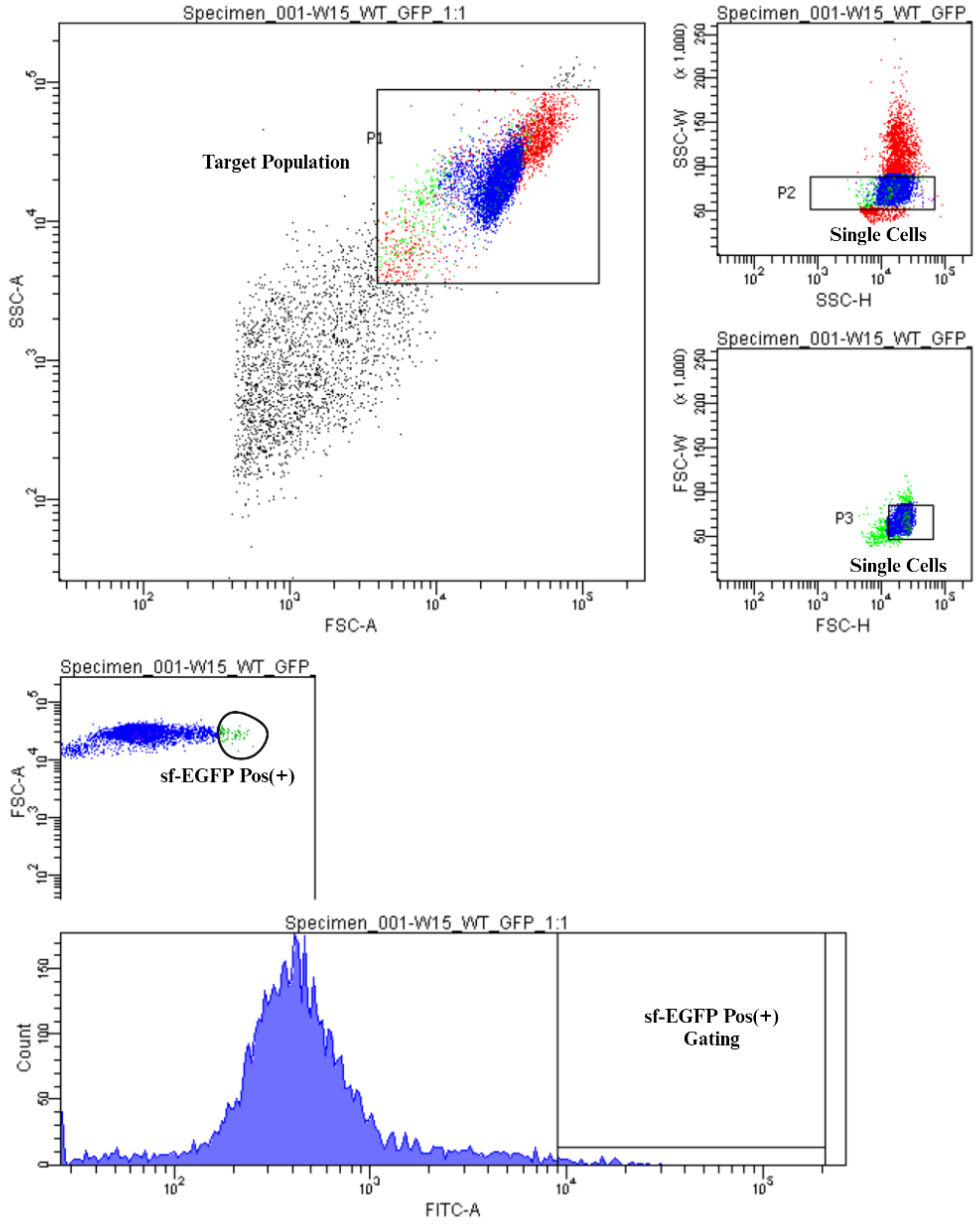


Figure S8. FACS sorting plots for PiggyBac clones generated using FACSDiva Version 6.1.3. The dot plots illustrate the selection process of representative WT PiggyBac plasmid in a 1:1 ratio with hyPBase transposase following sf-EGFP induction with 500 ng/ml dox. Starting from the top, the target population is selected (FSC-Area vs SSC-Area, log), and cells are subsequently gated for the single-cell population (SSC-Height vs SSC-Width and FSC-Height vs FSC-Width). From the single-cell population, the FITC-Area channel is used to select the sf-EGFP population, with further gating specified in the histogram (bottom).

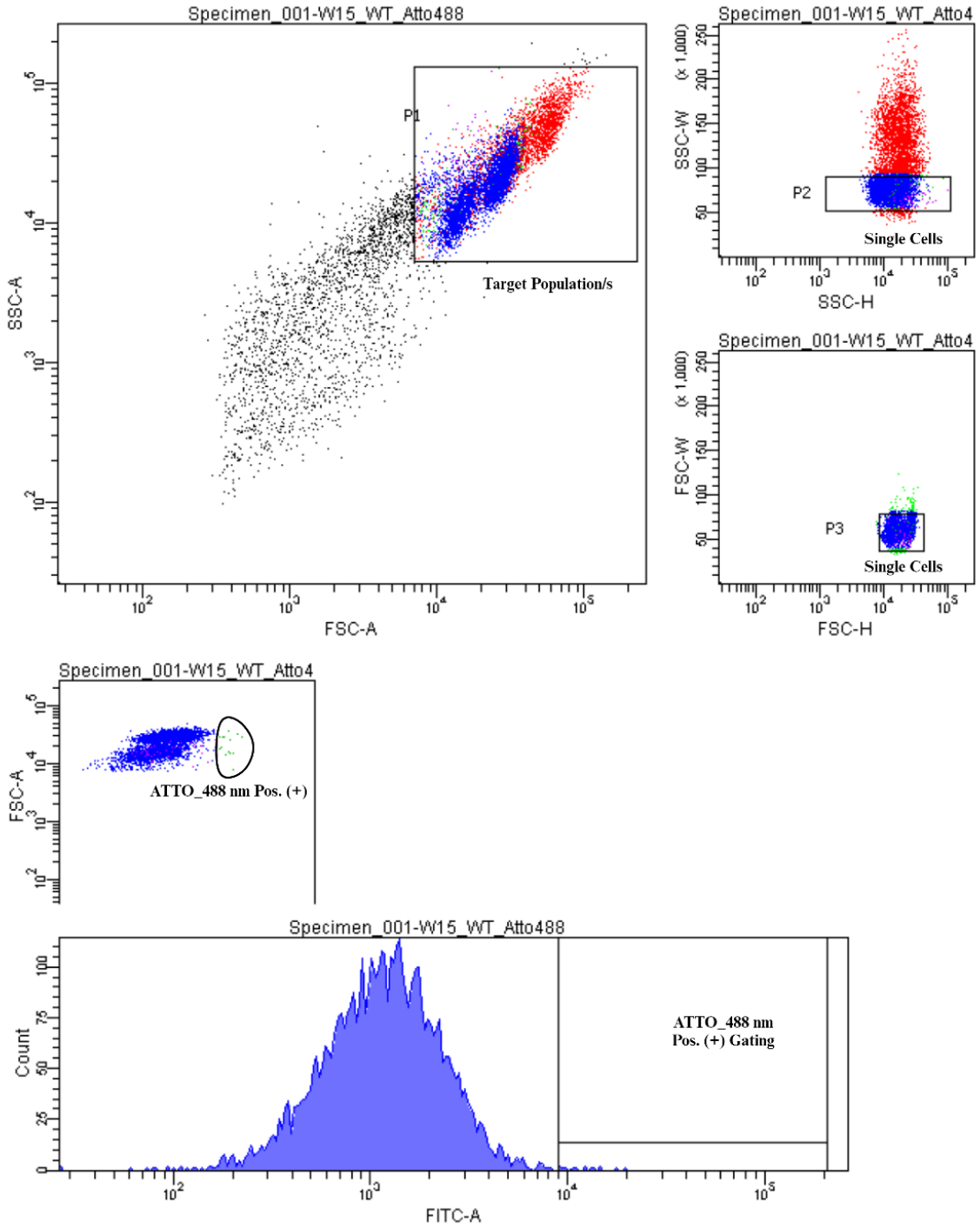


Figure S9. FACS sorting plots for CRISPR 3x-FLAG knock-in cell lines generated using FACSDiva Version 6.1.3. The dot plots illustrate the selection process of representative transfected W15 with crRNA and conjugated ATTO_488 nm tracer RNA, sorted 24 hrs post-transfection. Starting from the top, the target population is selected (FSC-Area vs SSC-Area, log), and cells are subsequently gated for the single-cell population (SSC-Height vs SSC-Width and FSC-Height vs FSC-Width). From the single-cell population, the FITC-Area channel is used to select the ATTO_488 nm tracer RNA positive population, with further gating specified in the histogram (bottom).

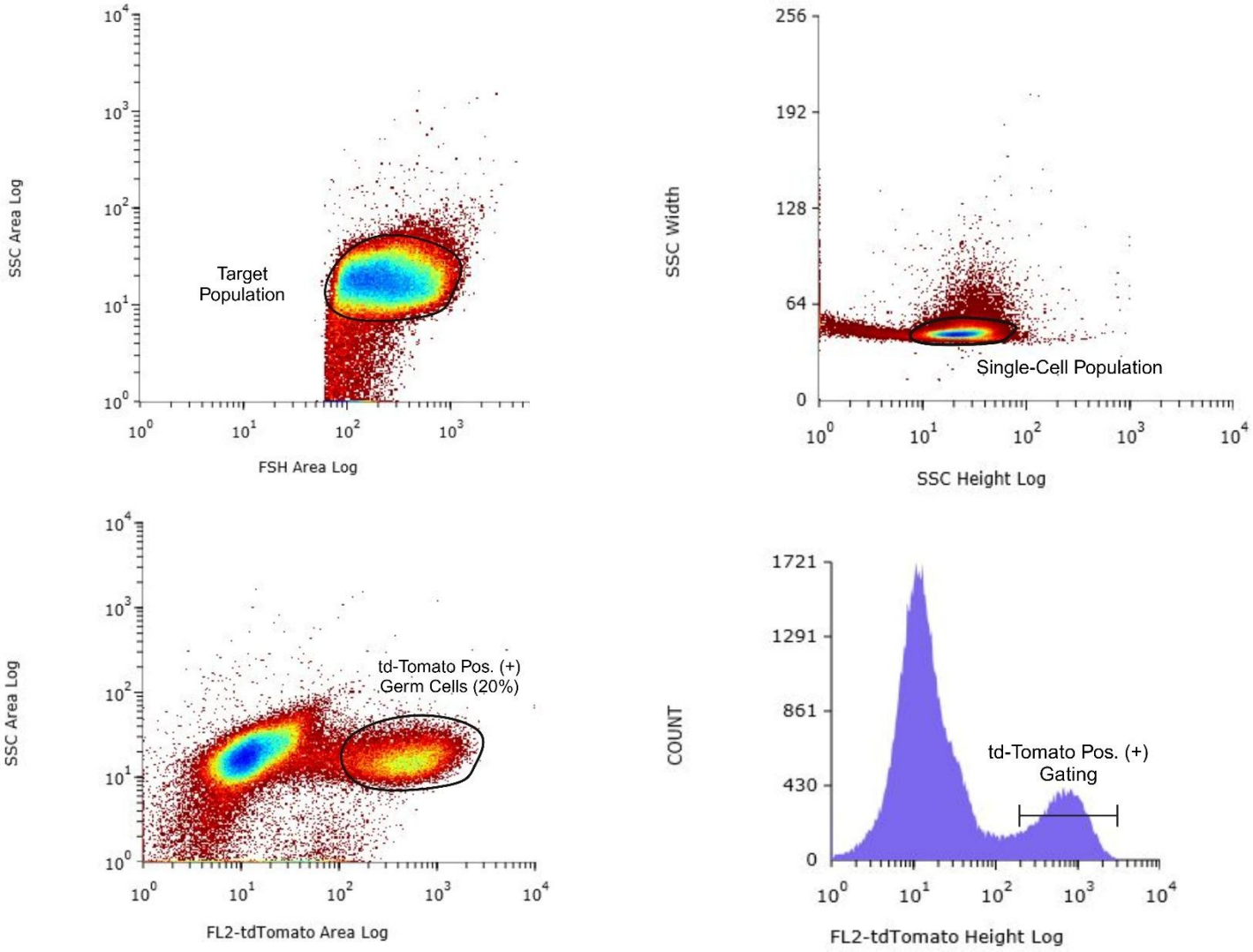


Figure S10. FACS workflow of unedited W15 cell line and cell lines uninduced by dox on Day 4 of EB differentiation, generated using ProSort™ Software. From the top-left, the target population is selected using an FSH Area Log vs SSC Area Log density plot, followed by gating for single cells based on SSC Height Log vs SSC Width parameters. From this population, cells are sorted for td-Tomato positive germ cells via the FL2-tdTomato Area Log channel vs SSC Area Log. The encircled population on the right represents the td-Tomato positive cells (20%), which are also gated in the histogram plot. The population on the left side is collected as the negative soma population.

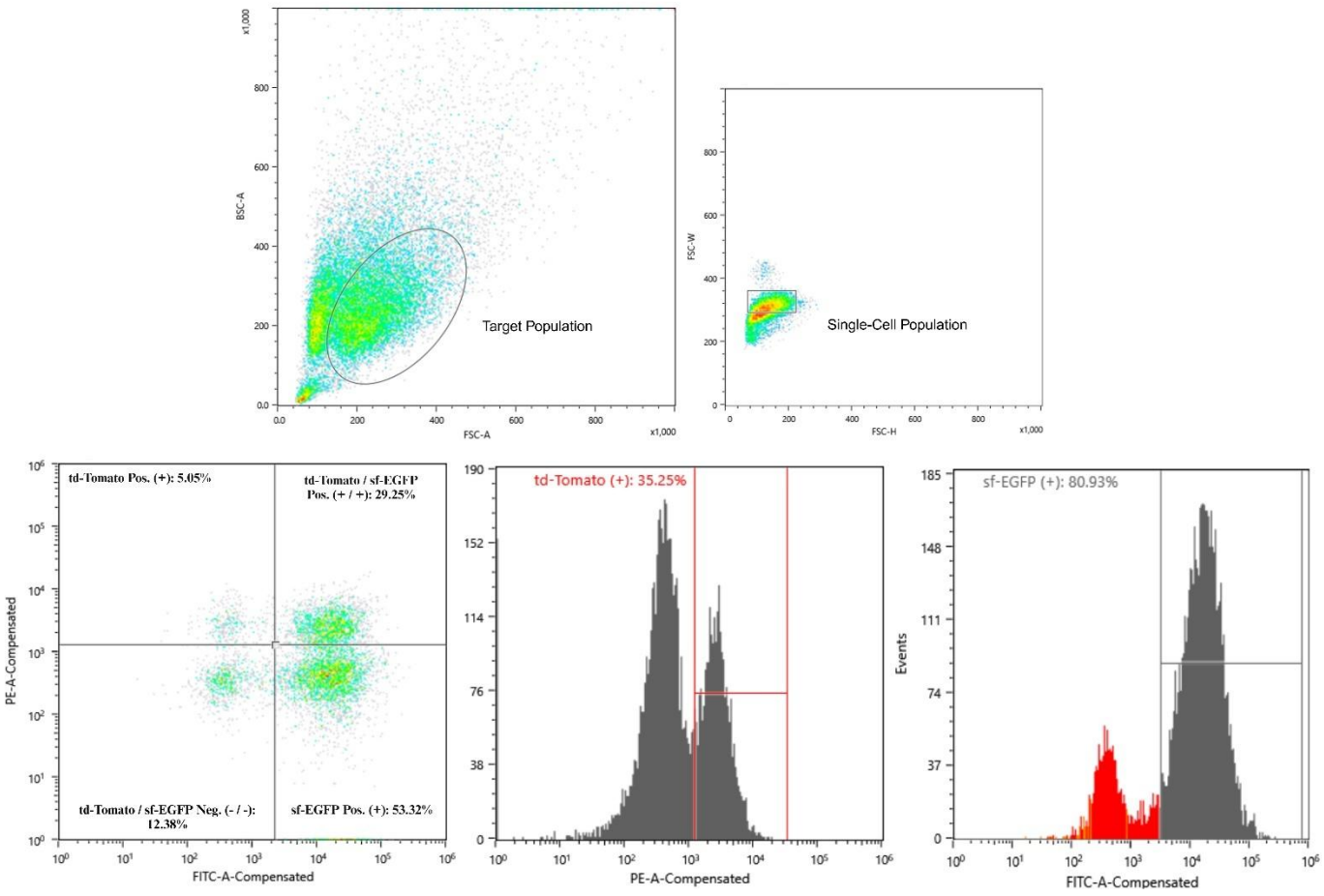


Figure S11. FACS workflow of representative edited PiggyBac cell lines induced with 500 ng/ml dox on Day 4 of EB differentiation, visualized using the Sonny MA900 Multi-Application Sorter Software. From the top left, the target population is selected using FSC-Area vs BSC-Area gating, followed by gating of the single-cell population via FSC-Height vs FSC-Width. From the bottom left, after automatic compensation of the FITC (sf-EGFP) and PE (td-Tomato) channels, the single-cell gated population is divided into four subpopulations through quadrant gating. For RNA sequencing, positive induced germ cells were sorted from the double-fluorescence marker subpopulation, while the negative population was sorted from the single gated sf-EGFP subpopulation. The histograms show the gating for each channel and fluorescence marker, guiding the gate setting of subpopulation quadrants.

RNA sample collection, quantification, library preparation RNA samples were collected from both established cell lines selected for analysis: WT 1.2.51 and MUT 3.2.5.2 cell lines, along with the unedited W15 control cell line. These samples were collected at various stages of differentiation, including stem cell, pre-me, Day 4, and soma FACS sorted populations (refer to **Fig. S10** and **Fig. S11** for FACS details). Each cell line at each differentiation stage was treated with and without dox and collected in three biological replicates for each condition (**Fig. S12**).

All samples were collected as pellets, digested in Trizol, and submerged in dry ice or liquid nitrogen. They were then homogenized for 5 min at RT and vortexed prior to adding chloroform (POCH, cat. 234431116). The samples were shaken vigorously for 15 sec, incubated at RT for 2-3 min, and centrifuged for 15 min at 12,000 x g at 4°C. This process resulted in a separation into a lower red phenol-chloroform

phase, an interphase, and a colorless upper aqueous phase. The aqueous phase was carefully transferred to a low-binding tube and mixed with an equal volume of ethanol (95-100%). The mixture was then transferred to a Zymo-Spin™ IC Column3 in a collection tube and centrifuged. RNA concentration and washing steps were performed according to the manufacturer's protocol and kit reagents (RNA Clean & Concentrator-5, Zymo Research, cat. R1014). The extracted RNA was quantified, and its integrity tested using the Agilent RNA 6000 Nano Kit (Agilent, cat. 5067-1511) on the Agilent 2100 Bioanalyzer system. A directional mRNA library (poly-A selection) with Illumina sequencing depth of 30M Paired-End (PE150) 150 bp length reads was prepared for 72 samples by Novogene and run on an Illumina HiSeq platform (25 Cambridge Science Park, Milton Road, Cambridge, CB4 0FW, United Kingdom). The total RNA from samples was approximately 400 ng, with RNA Integrity Numbers (RIN) ranging from 7 to 10. A summary of the library preparation, quality control, and sequencing is shown in **Fig. S13**. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. After fragmentation, the first strand cDNA was synthesized using random hexamer primers, followed by the second strand cDNA synthesis using dUTP instead of dTTP. The directional library was completed after end repair, A-tailing, Illumina adapter ligation, size selection, USER enzyme digestion, amplification, and purification. The library was checked for size distribution using Qubit, real-time PCR quantification, and a bioanalyzer. Quantified libraries were pooled and sequenced on Illumina platforms according to effective library concentration and data requirements.

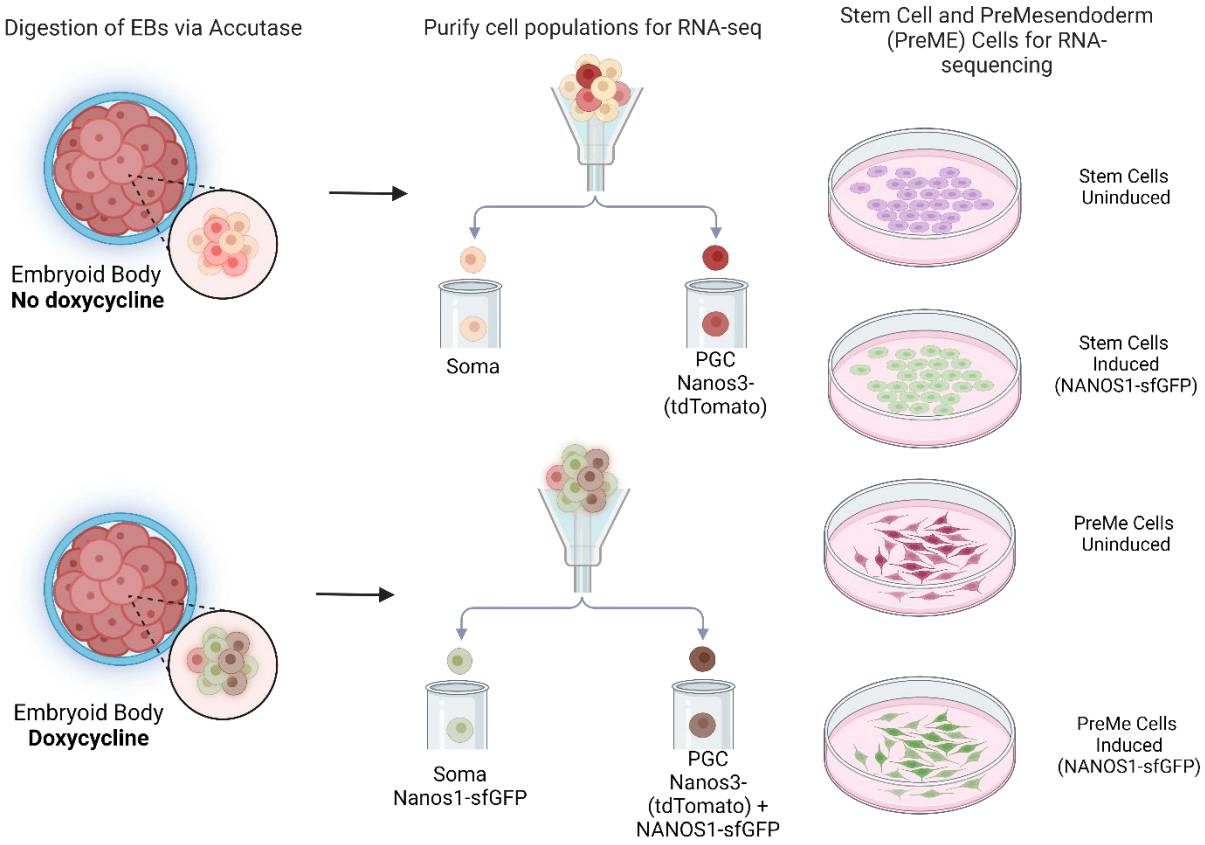


Figure S12. Schematic diagram illustrates the differentiation stages of PiggyBac clones: stem cell, pre-me, and Day 4 EBs. Cells were collected both before and after induction with dox (500 ng/ml). On Day 4 of differentiation, EBs were digested and sorted via FACS based on the td-Tomato germ cell marker and inducible sf-EGFP NANOS1 overexpression marker. At the pre-me and stem cell stages, cells were collected without sorting.

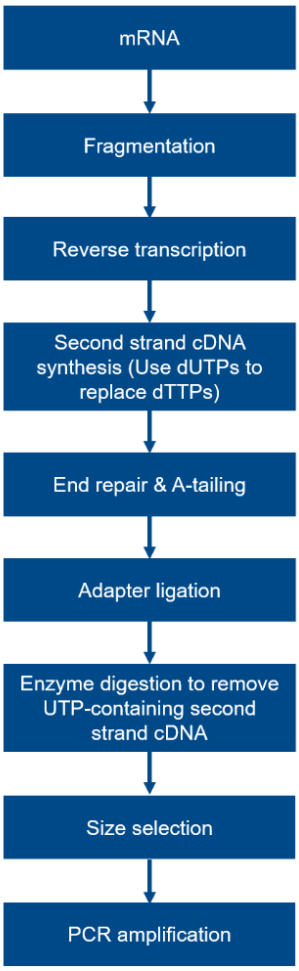


Figure S13. RNA library preparation workflow for bulk mRNA sequencing performed by Novogene. A total of 72 RNA samples were extracted from three cell lines at four stages differentiation: stem cells, pre-me, Day 4 germ cells, and Day 4 soma, both before and after dox induction of the NANOS1 cassette.

Immunofluorescence To analyze W15 and cell lines at both stem cell and pre-me stages, cells were cultured in monolayers on 8-well µ-Slides (ibidi, 80826). The cells were seeded at a density of 10,000 – 15,000 cells per well and initially cultured in 300 µl E8 media supplemented with ROCK Inhibitor. Monolayer cells underwent three washes with 200 µl 1x PBS for 10 min each. The cells were then fixed with 200 µl of 4% paraformaldehyde (Thermo Fisher Scientific, cat. A11313.22) in PBS at RT for 10 min, followed by three additional PBS washes. Subsequently, cells were permeabilized for 10 min with 0.25% Triton-X100 (Sigma-Aldrich, X100-100ML) in PBS. Next, samples were incubated with blocking buffer containing 0.1% Triton X-100, 5% normal donkey serum (Abcam, cat. ab7475), and 1% bovine serum albumin (BSA, Sigma-Aldrich, cat. A9647) in PBS. Primary antibodies, as specified in **Tab. S5**, were added at a final volume of 150 µl per well. For fluorescently tagged antibodies from the Human Pluripotent Stem Cell 3-Color Immunocytochemistry Kit (R&D Systems, cat. SC021) (see **Tab. S5** for SOX2, OCT3/4, and NANOG), the kit instructions were followed. Cells were incubated with all three conjugated antibodies in blocking buffer for 3 h at RT in the dark. For other non-conjugated antibodies, an overnight incubation at 4°C was performed,

except for the SSEA4 cytoplasmic marker. The following day, cells were washed three times with wash buffer (0.1% Triton X-100 in PBS) and incubated with Alexa fluorophore-tagged secondary antibodies (see **Tab. S5**), specific for the host species of the primary antibodies, in blocking buffer for 1 h at RT in the dark. Post incubation, the samples were counterstained with 1 µg/ml DAPI (Sigma-Aldrich, cat. F6057-20ML) in PBS for 10 min at RT, followed by three additional 10-min washes in PBS. Mounting medium (ibidi, cat. 50001) was then applied using a dropper (2-3 drops per well). The slides were visualized using the Leica Stellaris Confocal System (Leica DMI8 Microscope) and images were processed with the LasX software (LASX Office 1.4.6 284333, latest update March 22, 2024).

Table S5. List of antibodies used to validate cell lines via immunostaining during the stem cell and pre-me differentiation stages

Stage of Differentiation	Antibody	Supplier	Catalogue Number	Dilution (Final)	Incubation Length
Stem Cell	SOX2 NL557-Conjugated	R&D Systems	#967149 (part of kit SC021)	1:10	3 hours (4ºC)
Stem Cell	Oct-3/4 NL637-Conjugated	R&D Systems	#967150 (part of kit SC021)	1:10	3 hours (4ºC)
Stem Cell	NANOG NL493-Conjugated	R&D Systems	#967151 (part of kit SC021)	1:10	3 hours (4ºC)
Stem Cell	Human/Mouse SSEA-4 Antibody	R&D Systems	MAB1435-SP	1:160	3 hours (4ºC)
Stem Cell	Donkey anti Mouse IgG (H+L) Highly Cross Adsorbed Secondary Antibody, Alexa Fluor 488	ThermoFisher Scientific	A-21202	1:1000 Secondary Antibody against Human/Mouse SSEA-4 Primary Antibody	3 hours (4ºC)
PreMe	Human/Mouse Brachyury Antibody	R&D Systems	AF2085-SP	1:250	Overnight (4ºC)
PreMe	EOMES Recombinant Rabbit Monoclonal Antibody (BLR104H)	ThermoFisher Scientifc	MA5-44400	1:500	Overnight (4ºC)
PreMe	Alexa Fluor®647-conjugated AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch	715-605-151	1:800 Secondary Antibody against Human/Mouse Brachyury Primary Antibody	Overnight (4ºC)

PreMe	Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 647	ThermoFisher Scientific	A32795TR	1:500 Secondary Antibody against EOMES Rabbit Monoclonal Primary Antibody	Overnight (4ºC)

EB staining and light sheet imaging After generating EBs in 96-U bottom well plates at Day 4 of differentiation, organoids were collected in minimal medium and resuspended in 1 ml of 1% BSA (in PBS) on ice. The supernatant was discarded, and 1 ml of ice-cold 1x PBS added followed by cold 4% PFA in PBS retained in a 45-min incubation at 4°C. Following fixation, PFA was discarded, and EBs were washed with 1 mL PBST (0.1% Tween in PBS). Organoids were then incubated in organoid wash buffer (OWB: 2% BSA, 0.1% Triton X-100, 0.025% SDS in PBS) for 1 h at 4°C. EBs were transferred to a 48-well plate and incubated with primary antibodies (**Tab. S6**) diluted in OWB overnight with mixing at 4°C. After three washes with 300 µl OWB, secondary antibodies (**Tab. S6**) and DAPI (1:2000) in 100 µl OWB were added and incubated overnight with mixing at 4°C. Post-incubation, EBs were washed twice with PBST buffer and resuspended in 1x PBS at 4°C until imaging. For light sheet imaging using ZEISS Lattice Lightsheet 7, buffers including 0.1M PB Buffer, 0.02M Buffer, and Fruit Buffer (refractive index 1.475) were prepared (see **Tab. S7)**. The embedding gel was prepared by mixing 4% low gelling agarose in 0.02 M PB buffer with Fruit Buffer at 55-60°C to achieve a refractive index of 1.475. Organoids were incubated overnight in Fruit Buffer at RT, then transferred in minimal PBS to a 10 cm bacterial dish on a heating plate at 40°C and mixed with embedding gel using a capillary. The mixture was solidified in the fridge and visualized using Zeiss Lattice Lightsheet 7 with specific settings, processed via ZEN (black edition) version 3.1 LS.

Table S6. List of antibodies used to validate clonal cell lines via immunofluorescence on Day 4 of the EB differentiation stage

Stage of Differentiation	Antibody	Supplier	Catalogue Number	Dilution (Final)	Incubation Length
Day 4 EB	Human SOX17	R&D Systems	AF1924-SP	1:100	Overnight (4ºC)
Day 4 EB	PRDM14	Sigma-Aldrich	AB4350	1:100	Overnight (4ºC)
Day 4 EB	TFAP2C (AP-2γ: 6E4/4)	Santa Cruz Biotechnology	sc-12762	1:50	Overnight (4ºC)
Day 4 EB	Alexa Fluor®647-conjugated AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch	715-605-151	1:800 Secondary Antibody against TFAP2C (AP-2γ Antibody: 6E4/4)	Overnight (4ºC)
Day 4 EB	Donkey anti-Rabbit IgG (H+L) Highly	ThermoFisher Scientific	A32795TR	1:500	Overnight (4ºC)

	Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 647			Secondary Antibody against PRDM14 Antibody	
Day 4 EB	Alexa Fluor®647-conjugated AffiniPure Donkey Anti-Goat IgG (H+L)	Jackson ImmunoResearch	705-605-147	1:800 Secondary Antibody against Human SOX17 Antibody	Overnight (4ºC)

Table S7. List of buffers and reagents, specified by mass and volume, required for embedding EBs prior to light sheet imaging using the Zeiss Lattice Lightsheet 7 microscope

0.1M PB buffer	
Reagent	**Amount**
NaH_2PO_4_1x H_20	3.1 g
Na_2HPO_4	10.9 g
H_20	Up to 1l
0.1M PB buffer (Alternative)	
Reagent	**Amount**
NaH_2PO_4 (anhydrous)	2.7 g
Na_2HPO_4	10.9 g
H_20	Up to 1l
0.02M PB buffer	
Reagent	**Amount**
0.1M PB buffer	200 ml
H_20	up to 1l
Fruit buffer (refractive index 1.475) *	
Reagent	**Amount**
0.02 M PB buffer	17 ml
Fructose	40 g
Urea	16.5 g

*Mix reagents at 40ºC.

Karyotyping The W15 hESC line and PiggyBac cell lines were cultured to 70-80% confluency and then passaged into a 6-well plate per cell line. Colcemid (ThermoFisher Scientific, cat. 15212012) was added to the E8 culture medium at 10 µg/ml when numerous dividing cells were observed under the microscope. The colcemid treatment duration was optimized for each cell line, ranging from 1.5 to 2.5 h before collection. During collection, the culture media and PBS used for washing were retained to prevent loss of loosely attached cells. The mixture was centrifuged at 400 x g for 4.5 min, and a warmed hypotonic solution (37°C)

was added to the cells. The hypotonic solution consisted of 3 g KCl, 4.8 g 1M HEPES, 0.2 g EGTA, and water to a volume of 1L, pH 7.4. Dr n. med. Anna Przybyłowicz-Chalecka, (specjalista Laboratoryjnej Genetyki Medycznej) performed the subsequent steps of fixing, slide preparation, chromosome staining, G-banding analysis, and results interpretation (Karol Marcinkowski Medical University in Poznan, Hematology Diagnostics Laboratory).

Western Blot The overexpressed NANOS1-3xFLAG tagged cassette and endogenous NANOS1-3x FLAG were quantified via western blot analysis under standard conditions, using nitrocellulose or PVDF membranes and horseradish peroxidase (HRP)-conjugated secondary antibodies. Semi-quantitative and qualitative visualization of protein levels was performed using ImageLab 5.1 software (Bio-Rad, Hercules, California, USA). The chemiluminescent signal was detected using the Clarity™ Western ECL Substrate for HRP (Bio-Rad, cat. 1705060). The primary antibody commonly used was monoclonal anti-DDK (ANTI-FLAG® M2) [Sigma-Aldrich, cat. F1804, 1:1250 (endogenous); 1:1500 (overexpressed)] to detect the 3xFLAG tagged NANOS1 both endogenously and overexpressed. Anti-ACTB (actin beta) (Sigma-Aldrich, cat. A2066, 1:10000) and anti-VCL (vinculin) (Abcam, cat. ab129002, 1:20000) antibodies were used as reference controls. Secondary antibodies included goat anti-rabbit IgG-HRP (Sigma-Aldrich, cat. A6154, 1:25,000) and goat anti-mouse IgG-HRP (Santa Cruz Biotechnology, cat. sc-2005, 1:10,000).

Protein immunoprecipitation Cell pellets were collected from the pre-me stage of unedited and CRISPR-edited 3xFLAG-NANOS1 cell lines. After a 6 h pre-me period, the cells were scraped, centrifuged at 500 x g for 5 min, and resuspended in ice-cold 1x PBS. The cells were then transferred to a 1.5 ml low binding microcentrifuge tube, centrifuged, and the residual PBS was removed. The pellet weight was measured to determine the amount of precipitation beads required per pellet. The pellets were flash frozen in liquid nitrogen and stored at -80ºC. The precipitation protocol followed the instructions provided by the Dynabeads™ Co-Immunoprecipitation Kit (ThermoFisher Scientific, cat. 14321D). For this protocol, the primary monoclonal antibody anti-DDK (ANTI-FLAG® M2) (Sigma-Aldrich, F1804) and an equal amount of normal mouse IgG control (EMD Millipore, cat. 200621) were coupled to the surface of Dynabeads™ M-270 Epoxy beads overnight. The eluted samples (negative control, CRISPR-edited clone, and positive control) were then analyzed by western blot under standard conditions using a PVDF membrane and horseradish peroxidase (HRP)-conjugated secondary antibodies (refer to **Western Blot** section).

eCLIP (Enhanced Cross-Linking Precipitation) The assay was performed using the eCLIP Library Prep Kit (Eclipse Bioinnovations, cat. ECEK-0001) on 2 x 10^7 pre-me cultured WT-1.2.5.1 and MUT-NANOS1 3.2.5.2 cell lines before and after dox treatment in two biological replicates. The adherent cells were crosslinked using 254 nm UV-C Mercury Bulbs via the UVP crosslinker CL-3000 (Analytikjena, 849-95-0615-02). The monoclonal anti-FLAG M2 antibody (Sigma-Aldrich, F1804, 1000 µg /ml) was validated in house via the Antibody IP Validation Kit (Eclipse BioInnovations, cat. 143663) according to manufacturers' instructions. A summary diagram of the protocol from sample collection to library preparation is shown in

Fig. S14. The eCLIP methodology optimized by Eclipse BioInnovations (5770 Oberlin Dr. San Diego, USA) is described by Van Nostrand et al. (2016) and was applied with no further modifications (**https://eclipsebio.com/capabilities/rbp-eclip/** - last accessed 30/12/2024). Four input and four immunoprecipitated libraries were sequenced on an Illumina HiSeq platform using barcoding and a single-end (SE100) 100 bp length with a sequencing depth of 60M reads.

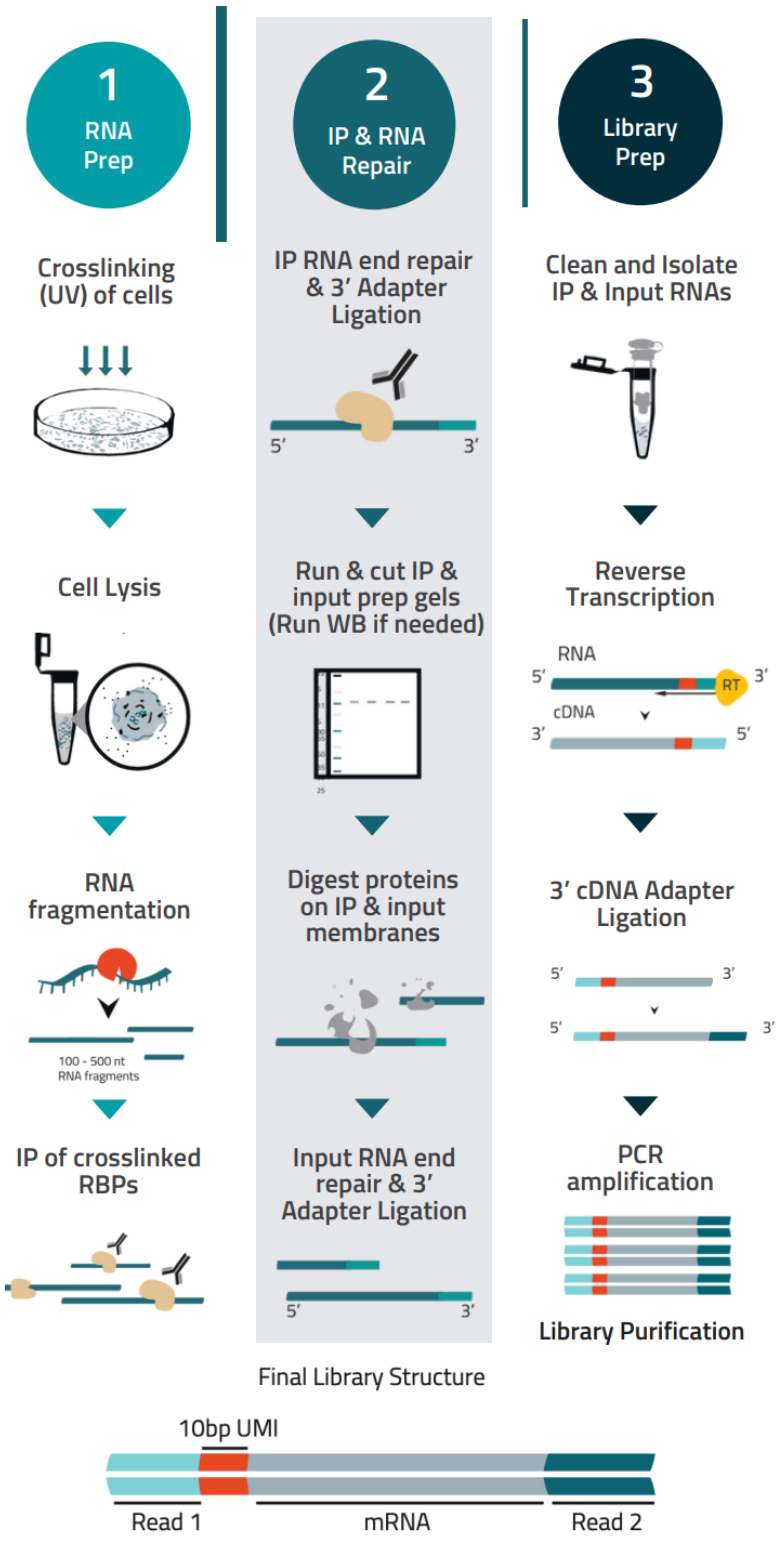


Figure S14. Summary of eCLIP assay from RNA preparation to sequencing library preparation (adapted from Eclipse Bioinnovations Datasheet). 1/ NANOS1 RNA Binding Protein (RBP)-RNA interactions are UV crosslinked and fragmented. 2/ The target NANOS1 is immunoprecipitated, a 3' RNA adapter is ligated, and the material is run on denaturing protein gels. 3/ It is then reverse transcribed to ssDNA with a second adapter, followed by PCR amplification for high-throughput sequencing.

Quantitative RT-PCR Total RNA was purified using Trizol extraction according to the manufacturer's instructions. cDNA synthesis was performed with 500 ng of RNA per sample using the QuantiTect Reverse Transcription Kit (Qiagen, cat. 205313). RT-qPCR was conducted with the PowerUp™ SYBR™ Green Master Mix for qPCR (ThermoFisher Scientific, cat. A25742) on the CFX Connect Real-Time PCR Detection System (Bio-Rad). The RT-qPCR was executed in technical triplicates per biological replicates (each sample was collected in 3 biological replicates) using Bio-Rad CFX Maestro software v2.2. Biological replicates were averaged to calculate mean fold changes, and results were expressed as fold differences compared to control samples, normalized against Alu repeats (Table S8) using the ΔΔCt method. Primer sequences are listed in **Table S8**.

Table S8. Primer sequences for target mRNAs exhibiting significant differential expression following MUT-NANOS1 overexpression. Notably, the majority of these genes were also enriched in the MUT-NANOS1 RNA interactome as identified by eCLIP analysis.

Gene	Forward Primer	Reverse Primer	Size (bp)
Alu repeats	CATGGTGAAACCCCGTCTCTA	GCCTCAGCCTCCCGAGTAG	N/A
AXIN2	GTCTCTACCTCATTTCCCGAGAAC	CGAGATCAGCTCAGCTGCA A	88
BAMBI	GCTGCTCACCAAAGGTGAAA	TTGCAAGAGAGTCCAGGCAG	161
BCL9L	GTACAGTGGGGACGAATGGT	ACATTACAATTGTGGGTGGCC	81
BMP4	GGAGCTTCCACCACGAAGAA	GGAAGCCCCTTTCCCAATCA	174
CCND1	ATCAAGTGTGACCCGGACTG	CTTGGGGTCCATGTTCTGCT	99
CD24	GGGAAGTGAAGACTGGAAGCTA	CAGTGTGTGACCATGCGAAC	279
CDH1	AGCCTGTCGAAGCAGGATTG	AGTCCTGGTCCTCTTCTCCG	204
CDH2	AGTCACCGTGGTCAAACCAA	ACAGACACGGTTGCAGTTGA	129
CTNNB1	GCTGGGACCTTGCATAACCT	TCCACTGGTGAACCAAGCAT	104
DACT1	GTCGGCCTAGCTCAGGGTTT	GACTCAAGGTCGCCTCCAAG	144
EOMES	CTGGCTTCCGTGCCCACGTC	CATGCGCCTGCCCTGTTTCG	101
FZD10	ACGAACGCCTCAACATGGAT	CGGTTTTCTCCGGCTCTTCT	251
HAND1	CCTACCTGATGGACGTGCTG	AGGAAAACCTTCGTGCTGCT	134
LEF1	GAACACCCCGATGACGGAAA	ATGTGTGACGGGTGTGATCC	257
LRP4	GCCGCCAAGTCATTATCT	TCAGCACCTTCCTCTTACT	144
LRP5	GAGACCGTACAGGCCCTACA	TGTAGTCGCTGTCACACACG	83
MYC1	GGACCCGCTTCTCTGAAAGG	TAACGTTGAGGGGCATCGTC	104

NANOG	TGCTGAGATGCCTCACACGGA	TGACCGGGACCTTGTCTTCCTT	155
NANOS3	TTGACCTGTGGACAGATTAC	TGGTTTCAGGACCCTCTT	75
OCT4 (POU5F1)	GCTGGAGCAAAACCCGGAGG	TCGGCCTGTGTATATCCCAGGGTG	114
PRDM1	GCGACTCCATTCTGGAGAGA	GTGCACAAACTGGGTGAACT	70
PRDM14	TCTGGAGACAGACCATACCAGT	CCTGATGTGTGTGCGGAGTA	78
SDC1	GGAGCTCCACTCTGCTTCTC	GGGAGAGGCTGCTTCAGTTT	140
SFRP2	GTGGCTCAAAGACAGCTTGC	CTCCCCACCCTGTTTCTGTC	91
SOX17	CAAGGGCGAGTCCCGTATC	ACGACTTGCCCAGCATCTTG	131
TBX3	TTCCTGCACCAACACCAATA	ACTTTCCCCCACTTTTGGAC	86
TCF4	TCTGAATCCGAAAGCTGCGT	TGTGGCAACTTGGACCCTTT	149
TCF7	CCAAGAATCCACCACAGGAGG	CGCAGGGCTAGTAAGCAGTT	183
TFAP2C	AAAGCCGCTCATGTGACTCT	AGGAAATTCGGCTTCACAGA	96
TCF7L1	GATCCAAAGACAGGAATCCCC	GAGGAAGGGAGTACATGGGC	145
THBS1	GAGTTCAGGGCTCCTGTCG	CAGCAGGGATCCTGTGTGT	136
VCL	CCAGTGTTTCATACGCGCAC	TCTCTTTTCCAACCCGGACG	178
WLS	GTTTCCCTGGCTTACCGTGA	CATGCTCTGGAGTCTTGGGA	112
WNT5A	GAAGCCAATTCTTGGTGGTCG	TGCAGTTCCACCTTCGATGT	223
WNT5B	GCCAACTCCTGGTGGTCATT	TGGCATTCCTTGATGCCAGT	185

Wnt Inhibition Wnt-C59 (Tocris, cat. 5148) was dissolved in DMSO to prepare a 10 mM stock solution. A final working concentration of 10 μM was used and added on Day 3 of EB differentiation of the MUT 3.2.5.2 NANOS1 cell line along with 500 ng/ml dox. Cells were sorted as shown in **Fig. S11** and pellets were collected. RNA was isolated, converted to cDNA, and RT-qPCR was performed on both samples and controls for germ cell markers (NANOG, OCT4, PRDM14, TFAP2C, and NANOS3) as well as mesodermal markers (BMP4, HAND1, LEF1, TBX3, and CTNNB1). Primer sequences are provided in **Table S8**.

Clonal generation via puromycin selection and cloning The first set of PiggyBac plasmids featuring a puromycin selection cassette and m-Venus fluorescence marker (**Fig. S3A-B**) was used for cell line generation via puromycin selection and cloning cylinders (Pyrex® cloning cylinder, Sigma-Aldrich, CLS31666-125EA).

For this initial trial of transfection, a 1:1 ratio of transposase to transposon was applied. The seeding, plating, and selection of PiggyBac clones followed the protocol established by Han et al. (2022). After electroporation, cells were plated in 6-well plates (**Fig. S14**) and cultured in media containing 500 ng/ml puromycin (ThermoFisher Scientific, cat. A1113802) 48 h post-electroporation. An untransfected well (negative control) was used to assess the efficiency of the selection agent. Within 3 days, the untransfected cells died, and puromycin treatment was continued to maintain positive selection pressure on the edited cells. The surviving cells were re-seeded at low density into a 10-cm culture dish to allow single-cell colony growth. Individual colonies were marked from the bottom of the plate without removing the medium (**Fig. S14**), avoiding those in close proximity to each other. Sterilized cloning cylinders, coated with high-vacuum silicone grease (Sigma-Aldrich, cat. Z273554-1EA) to prevent spillage, were placed on marked colonies. 100 µl of trypsin was added to the cylinders for 1 min, and the trypsinized cell suspension was collected with a 200 µl micropipette and transferred to a 6-well plate. Once confluent, cells were frozen and characterized by flow cytometry for the presence of the m-Venus marker. Further enrichment of m-Venus's expression was performed via FACS to create more homogeneous cell lines.

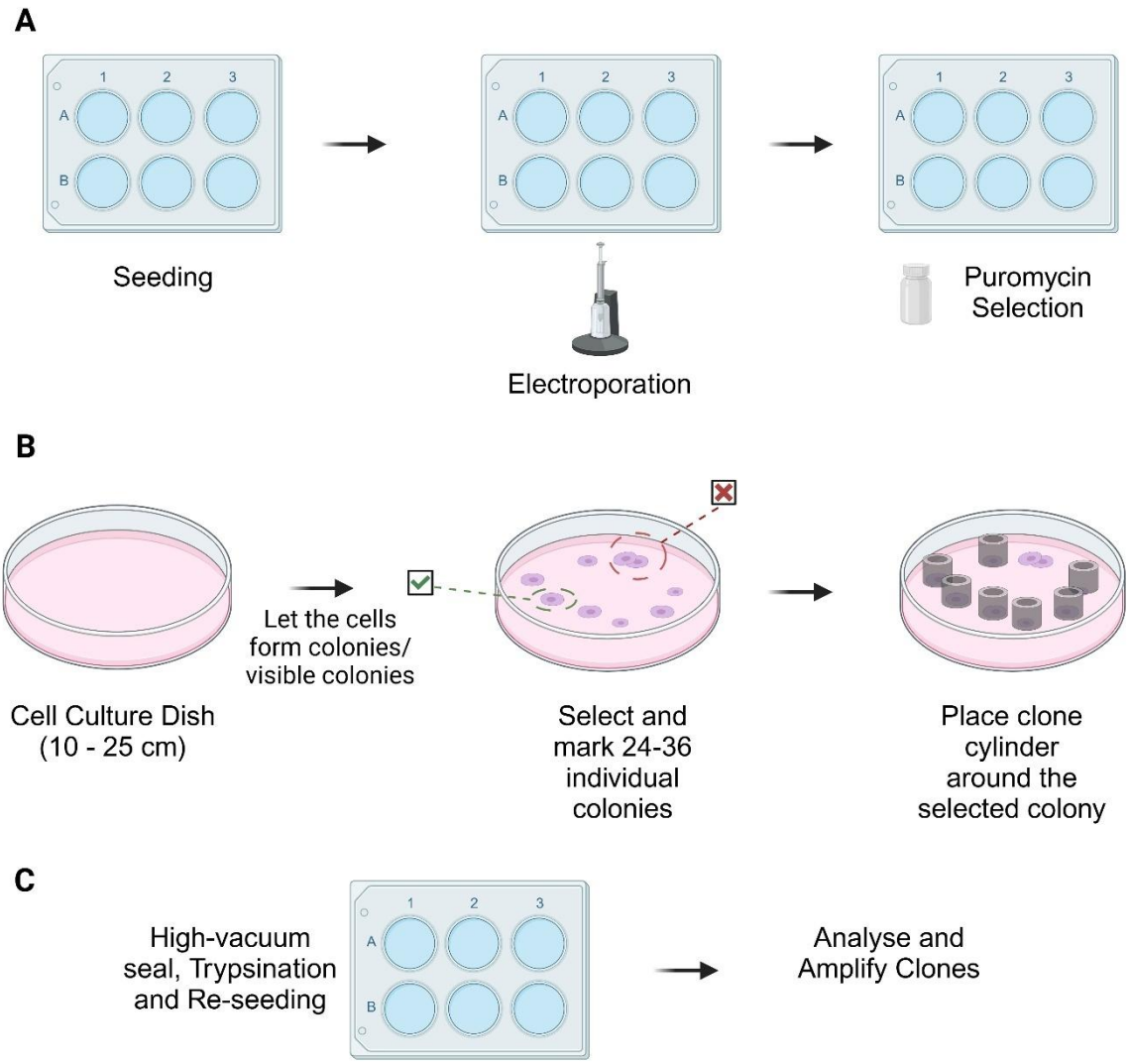


Figure S15. Schematic diagram of generating PiggyBac clones via antibiotic selection and cloning cylinders. A- Culture, seeding of cells post-electroporation in 6-well plate and selection of edited cells via puromycin for 1-week. **B-** Culturing selected cells in 10 – 25 cm cell culture dish at low density to obtain single-cell colonies. Properly isolated colonies were marked for sealing with cloning cylinders. **C-** Within the vacuum sealed cylinder, cells were trypsinized and moved to a 6-well plate for re-seeding and further culture (1-well per colony). Cell lines were further characterized post-storage in liquid nitrogen.

Analysis of copy number for the NANOS1 transgene cassette in PiggyBac cell lines To assess how transgene copy number influences the expression of target genes, the copy numbers were quantified using qPCR. An absolute quantitative standard curve was created by plotting ΔCt ($\Delta Ct = Ct_{TRE3G} - Ct_{GAPDH}$) against the log of TRE3G gene copies from extracted plasmid DNA from both WT and MUT cell lines (**Fig. S4A-B**) prepared in a ten-fold serial dilution series with at least five points (see **Tab. S9**). The resulting nanograms of diluted amplicon were converted to copy number using the formula:

$$\frac{X\ \text{ng} * 6.0221 \times 10^{23}\ \text{molecules/mol}}{(N * 660\ \text{g/mol}) * 1 \times 10^{9}\ \text{ng/g}}$$

where X is the amount of amplicon (ng), N is the length of the dsDNA amplicon, 660 g/mol is the average mass of 1 bp dsDNA, 6.022 x 10^{23} is Avogadro's constant, and 1 x 10^{9} is a conversion factor. These conversions were automated via the https://scienceprimer.com/copy-number-calculator-for-realtime-pcr based on the formula.

The standard curves were calculated as follows: for WT cell lines, log2N (copy number) = −0.6901ΔCt + 2.5182 (R^2=1), and for MUT cell lines, log2N = -0.7341ΔCt + 2.8797 (R^2=1) (**Fig. S16A-B**). Using these curves, the copy numbers of the transgene copies were determined (refer to **Fig. 4A-B**). The Ct values reported were averaged from triplicates. The primers used for TRE3G were Forward: GCTTTAGGCGTGTACGGTGG and Reverse: TGTGGAATTGC-TCCAGGCGA, with a product length of 79 bp, while for GAPDH, the primers were Forward: CTTCAACAGCGACACCCACT and Reverse: AGAGTTGTCAGGGCCCTTTTTCT, with a product length of 175 bp.

Table S9. Standard curve conversion table to mediate plasmid copy numbers per cell line. The WT and MUT PiggyBac plasmids (**Fig. S4A-B**) were serially diluted ten-fold and converted to dsDNA target copies. These values were then correlated with the Ct values of TRE3G (Tet-On promoter for both cassettes) compared to GAPDH (housekeeping gene).

WT PiggyBac Plasmid Cassette				
dsTarget copies/ul	Ct (TRE3G)	Ct (GAPDH)	ΔCT (t)	Log10 Target Copies
98,800,000	16.03117615	23.2614059	-7.230	7.994756945
988,000	21.51506754	23.1336655	-1.619	5.994756945
9880	22.97868134	23.2171249	-0.238	3.994756945
98.8	22.96800146	23.11147	-0.143	1.994756945
0.988	23.52448021	23.1312872	0.393	-0.005243055
MUT PiggyBac Plasmid Cassette				
dsTarget copies/ul	Ct (TRE3G)	Ct (GAPDH)	ΔCT (t)	Log10 Target Copies
98,800000	15.86780871	23.1676335	-7.300	7.994756945
988000	21.57993453	23.1010698	-1.521	5.994756945
9880	23.01647911	23.0693915	-0.053	3.994756945
98.8	23.66189177	23.0822262	0.580	1.994756945
0.988	23.67592056	23.071975	0.604	-0.005243055

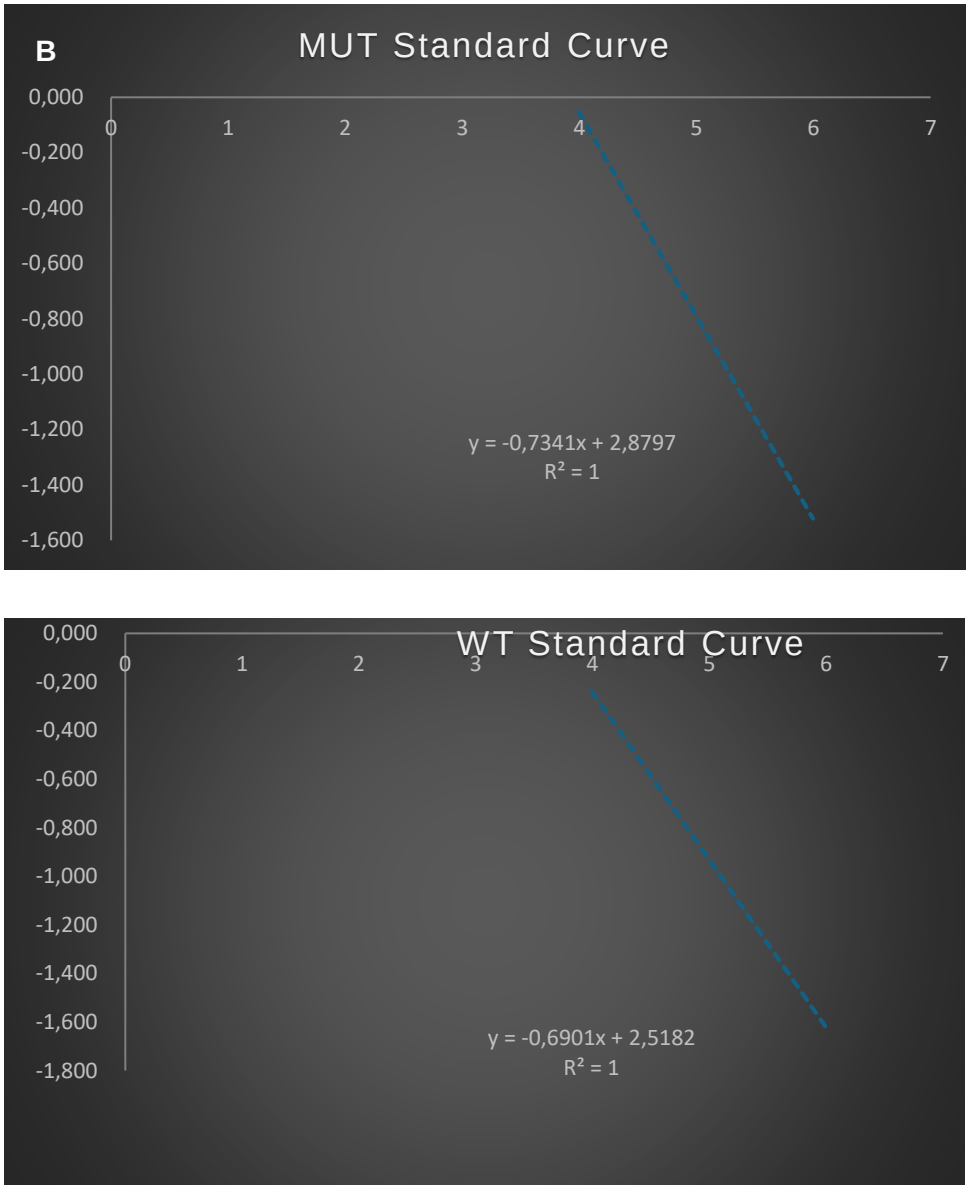


Figure S16. Establishment of the absolute quantitative standard curve to determine PiggyBac plasmid copy numbers per clone generated. The standard samples used to create the **A-** WT and **B-** MUT standard curves are detailed in **Table S9**. The absolute quantitative standard curve was generated by plotting ΔCt ($\Delta Ct = Ct_{TRE3G} - Ct_{GAPDH}$) against the log of TRE3G gene copies from the corresponding standard samples.

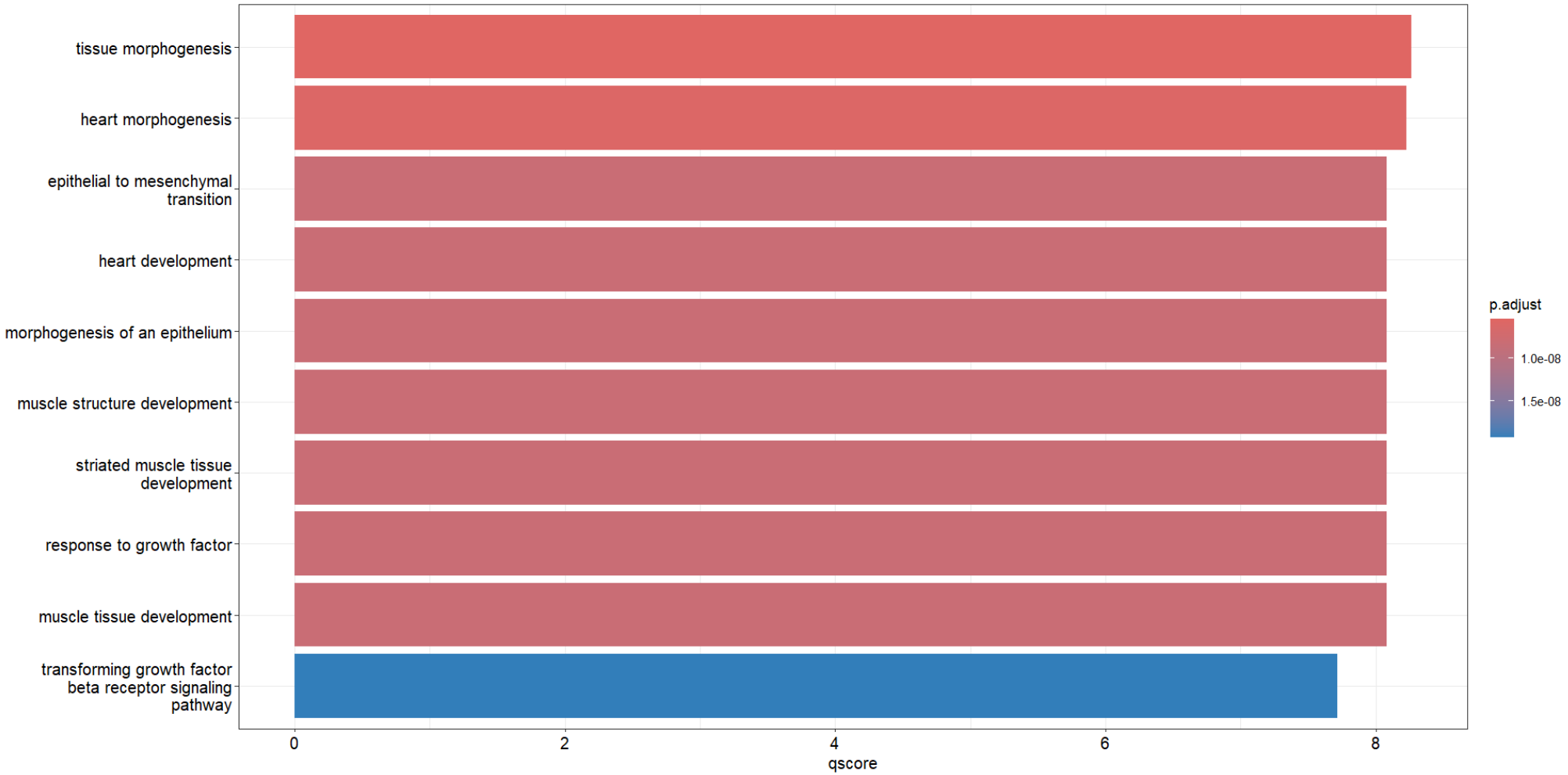

tissue morphogenesis
heart morphogenesis
epithelial to mesenchymal transition
heart development
morphogenesis of an epithelium
muscle structure development
striated muscle tissue development
response to growth factor
muscle tissue development
transforming growth factor beta receptor signaling pathway
0
2
4
6
8
qscore
p.adjust
1.0e-08
1.5e-08

Figure S17. Top 10 upregulated pathways from Gene Ontology (GO) analysis comparing primordial germ cell-like cells (PGCLCs) overexpressing WT- versus MUT-NANOS1. Upregulated pathways in MUT PGCLCs include tissue morphogenesis, heart development, epithelial-to-mesenchymal transition, and transforming growth factor beta receptor signaling mediating differentiation and mesoderm-like differentiation trajectory compared to WT PGCLCs (DESeq analysis). The x-axis lists enriched biological pathways, while the y-axis represents their significance (qscore). Bar colors indicate adjusted p-values (padjust), with red showing higher significance.

Bioinformatic analysis of eCLIP data eCLIP data were processed using the Skipper pipeline (https://github.com/YeoLab/skipper, DOI: 10.1016/j.xgen.2023.100317). Two experiments were analyzed, one for WT- and one for MUT-NANOS1, each comprising two eCLIP replicates and two size-matched input (SMI) controls. Initially, the human genome reference (GRCh38 primary assembly) and its corresponding GFF3 annotation file were downloaded from GENCODE version 38. The parse_gff.R script was then used to tile windows across annotated transcripts, generating fixed bins to efficiently aggregate read start signals across samples. Read 1 was designated as the informative read containing the mRNA sequence, and the selected input reads were employed to model the variance in read counts. The main Skipper pipeline was subsequently executed with a modification: UMI-tools was used for UMI removal in place of the default umicollapse. Software versions and R packages used within the Skipper pipeline are listed in **Table S10** and **Table S11**.

Table S10. List of software versions employed in the Skipper pipeline as part of eCLIP data analysis

Software	Version	Link	Reference (doi)
skewer	0.2.2	https://github.com/relipmoc/skewer	10.1186/1471-2105-15-182
STAR	2.7.10b	https://github.com/alexdobin/STAR	10.1093/bioinformatics/bts635
UMI-tools	0.4.4	https://github.com/CGATOxford/UMI-tools	10.1101/gr.209601.116
samtools	1.5	https://github.com/samtools/samtools	10.1093/gigascience/giab008
bedtools	2.31.0	https://github.com/arq5x/bedtools2	10.1093/bioinformatics/btq033
R	4.2.3	https://www.R-project.org/	10.59350/t79xt-tf203
HOMER	4.11	http://homer.ucsd.edu/homer/	10.1016/j.molcel.2010.05.004
awk	5.1.0	https://www.gnu.org/software/gawk/	https://www.gnu.org/software/gawk/manual/gawk.html

Table S11. List of R package versions utilized in the Skipper pipeline as part of eCLIP data analysis

R package	Version
_r-mutex	1.0.1
r-askpass	1.1
r-assertthat	0.2.1
r-backports	1.4.1
r-base64enc	0.1_3
r-bit	4.0.5
r-bit64	4.0.5
r-blob	1.2.4
r-boot	1.3_28.1
r-broom	1.0.5
r-bslib	0.5.1
r-cachem	1.0.8
r-callr	3.7.3
r-cellranger	1.1.0
r-class	7.3_22
r-cli	3.6.1
r-clipr	0.8.0
r-cluster	2.1.4
r-codetools	0.2_19
r-colorspace	2.1_0
r-conflicted	1.2.0
r-cpp11	0.4.6
r-crayon	1.5.2
r-curl	4.3.3
r-data.table	1.14.8
r-dbi	1.1.3

r-dbplyr	2.3.3
r-digest	0.6.33
r-dplyr	1.1.2
r-dtplyr	1.3.1
r-ellipsis	0.3.2
r-evaluate	0.21
r-fansi	1.0.4
r-farver	2.1.1
r-fastmap	1.1.1
r-fontawesome	0.5.2
r-forcats	1.0.0
r-foreign	0.8_84
r-fs	1.6.3
r-gargle	1.5.2
r-generics	0.1.3
r-ggplot2	3.4.3
r-ggrepel	0.9.3
r-glue	1.6.2
r-googledrive	2.1.1
r-googlesheets4	1.1.1
r-gridextra	2.3
r-gtable	0.3.4
r-haven	2.5.3
r-highr	0.1
r-hms	1.1.3
r-htmltools	0.5.6
r-httr	1.4.7
r-ids	1.0.1

r-isoband	0.2.7
r-jquerylib	0.1.4
r-jsonlite	1.8.7
r-kernsmooth	2.23_22
r-knitr	1.43
r-labeling	0.4.3
r-lattice	0.21_8
r-lifecycle	1.0.3
r-lubridate	1.9.2
r-magrittr	2.0.3
r-mass	7.3_60
r-matrix	1.6_1
r-memoise	2.0.1
r-mgcv	1.9_0
r-mime	0.12
r-modelr	0.1.11
r-munsell	0.5.0
r-nlme	3.1_163
r-nnet	7.3_19
r-openssl	2.0.5
r-pillar	1.9.0
r-pkgconfig	2.0.3
r-prettyunits	1.1.1
r-processx	3.8.2
r-progress	1.2.2
r-ps	1.7.5
r-purrr	1.0.1
r-r6	2.5.1

r-ragg	1.2.5
r-rappdirs	0.3.3
r-rcolorbrewer	1.1_3
r-rcpp	1.0.11
r-readr	2.1.4
r-readxl	1.4.3
r-recommended	4.2
r-rematch	1.0.1
r-rematch2	2.1.2
r-reprex	2.0.2
r-rlang	1.1.1
r-rmarkdown	2.24
r-rpart	4.1.19
r-rstudioapi	0.15.0
r-rvest	1.0.3
r-sass	0.4.7
r-scales	1.2.1
r-selectr	0.4_2
r-spatial	7.3_17
r-stringi	1.7.12
r-stringr	1.5.0
r-survival	3.5_7
r-sys	3.4.2
r-systemfonts	1.0.4
r-textshaping	0.3.6
r-tibble	3.2.1
r-tidyr	1.3.0
r-tidyselect	1.2.0

r-tidyverse	2.0.0
r-timechange	0.2.0
r-tinytex	0.46
r-tzdb	0.4.0
r-utf8	1.2.3
r-uuid	1.1_1
r-vctrs	0.6.3
r-vgam	1.1_6
r-viridis	0.6.4
r-viridislite	0.4.2
r-vroom	1.6.3
r-withr	2.5.0
r-xfun	0.4
r-xml2	1.3.5
r-yaml	2.3.7

In this study, the binding sites of WT-NANOS1 and MUT-NANOS1 were compared to those of other RNA-binding proteins (RBPs) obtained by eCLIP and available through the ENCODE project (https://www.encodeproject.org/). A t-distributed stochastic neighbour embedding (t-SNE) algorithm was employed to generate a two-dimensional projection of the multidimensional data. The t-SNE data from ENCODE eCLIP experiments were obtained using the skipper pipeline, and the resulting plot was generated with the consult_encode_reference.R script from the same pipeline. This script utilizes the R-tsne library (https://github.com/jkrijthe/Rtsne; http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf; https://jmlr.org/papers/volume15/vandermaaten14a/vandermaaten14a.pdf) to compute the t-SNE data and produce the corresponding plots.

To determine whether the binding sites of WT- and MUT-NANOS1 were significantly different, a contingency table was constructed that catalogued the number of reproducible enriched windows in each category, as defined by the “feature_type_top” column, and a chi-square test was applied to this table. A similar analysis was performed for length-normalized binding sites. First, the reproducible enriched windows mapping to 5’ untranslated regions (5’UTRs), coding sequences (CDSs), introns, and 3’ untranslated regions (3’UTRs) were filtered, and the total lengths of these regions were calculated using the Gencode v.38 GTF

file in conjunction with the GenomicFeatures R library (https://github.com/Bioconductor/GenomicFeatures; doi:10.1371/journal.pcbi.1003118). Normalization factors were then established using the 5'UTRs as the reference, with the total lengths of the other regions divided by the total length of the 5'UTRs. The total lengths and corresponding normalization factors are presented in **Tab. S12**. Finally, the number of reproducible enriched windows for a given region was multiplied by the corresponding normalization factor, and a contingency table was prepared as described above. A chi-square test was subsequently applied to this analysis.

Table S12. Length-normalized target RNA binding sites, which list the total nucleotide lengths for different RNA regions along with their normalization factors. These factors are derived by comparing each region's length to that of the 5'UTR (set as the reference with a factor of 1) and are critical for adjusting the number of reproducible enriched windows across the various RNA regions. This adjustment enables a more accurate comparison of binding site distributions between the WT- and MUT-NANOS1.

Target RNA region	Total length (nt)	Normalization factor
5'UTR	20594626	1
CDS	44048197	0.467547537
Introns	9089276557	0.002265816
3'UTR	99884601	0.206184194

Motif enrichment analysis Motif enrichment analysis was performed using DREME (https://meme-suite.org/meme/doc/dreme.html, Bailey et al., 2009). Sequences from enriched windows identified by the Skipper pipeline were compared against a background generated by shuffling these sequences while preserving dinucleotide frequencies, a standard procedure for background definition. Motifs were identified for WT- and MUT-NANOS1, using a threshold of occurrence in at least 10% of the enriched windows (approximately 370 for WT and 420 for MUT) and an E-value cutoff of ≤ 0.05, where the E-value was calculated as the product of the p-value and the number of motif occurrences in the positive set. Motif occurrences were plotted across genomic features (e.g., introns, CDS, 3'UTRs).

Bioinformatic analysis of RNA sequencing data Transcript quantification was performed using Salmon v1.10.2 (https://github.com/COMBINE-lab/salmon, doi: 10.1038/nmeth.4197) with the human reference transcriptome (GRCh38) obtained from Gencode v38. Initially, a Salmon index was generated from the combined genome and transcriptome ("gentrome") as described in the Salmon documentation (https://salmon.readthedocs.io/en/latest/salmon.html#preparing-transcriptome-indices-mapping-based-mode). Raw reads were first adapter-trimmed using Cutadapt v4.4 (https://cutadapt.readthedocs.io/, doi:10.14806/ej.17.1.200) before quantification with the Salmon quant command. The resulting output was imported into R v4.1.3 using the tximport library v1.22.0

(https://bioconductor.org/packages/release/bioc/html/tximport.html, doi:10.12688/f1000research.7563.2). Differential gene expression analysis was then conducted using either DESeq2 v1.34.0

(https://bioconductor.org/packages/release/bioc/html/DESeq2.html, doi:10.1186/s13059-014-0550-8) or ImpulseDE2 v1.8.0

(https://bioconductor.riken.jp/packages/3.9/bioc/html/ImpulseDE2.html, doi:10.1093/nar/gky675). DESeq2 was employed to assess differential expression between two conditions, while ImpulseDE2 was used to analyze differential expression across pseudotime points, where each developmental stage was treated as a discrete time point to capture transient gene regulation. Differential expression analyses were performed separately for WT-NANOS1 and MUT-NANOS1 samples as well as control W15 samples, with dox-treated samples designated as "target case samples" and untreated samples as "controls." Additionally, a comparison between WT- and MUT-NANOS1 dox-treated samples was performed, designating the MUT-NANOS1 samples as the “target case” and the WT-NANOS1 samples as “the controls”. Furthermore, comparisons were also made between both WT- and MUT-NANOS1 samples and the W15 dox-treated samples, with W15 serving as the control in these analyses. In all cases, pseudotime points were assigned as follows: stem cells → 1, preme → 2, PGCLs → 3, and soma → 4. Gene blacklisting was applied during differential expression analysis with DESeq to exclude genes whose expression was affected by dox treatment, as determined from comparisons within the W15 control cell line before and after treatment. Additionally, genes altered by inherent differences between cell lines prior to dox treatment were also blacklisted. This approach effectively removed confounding factors associated with dox administration and cell line generation, thereby refining the assessment of NANOS1 variant-specific effects.

Data Visualization All downstream analyses were performed in R version 4.1.3. Data visualizations were generated using the ggpubr (Kassambara, 2022) and ggplot2 (Wickham, 2016) packages to create publication-ready plots. Venn diagrams were constructed with ggvenn (https://github.com/yanlinlin82/ggvenn; doi:10.32614/CRAN.package.ggvenn). Heatmaps and MA plots were generated in R using several libraries, including heatmap (v1.0.12) and ggplot2 (v3.5.1). Gene annotations were imported from the GENCODE v38 GTF file and converted into a data frame to map ENSEMBL IDs

(https://www.ensembl.org/Homo_sapiens/Info/Index) to gene names. A DESeq2 object (dds) was then loaded from an RDS file containing differential expression results. Normalized count data were extracted from the DESeq2 object, annotated similarly, and used to generate a heatmap with row-wise scaling and Euclidean distance clustering applied to both rows and columns (using the pheatmap function, https://cran.r-project.org/web/packages/pheatmap/index.html). For the MA plot, the DESeq2 results were converted to a data frame and visualized using the plotMA function (Love et al., 2014; https://search.r-project.org/CRAN/refmans/bulkAnalyseR/html/ma_plot.html), with the y-axis limits set from -5 to 5 to clearly display the relationship between $\log_2$ fold-change and mean expression levels.

Functional Analysis All downstream functional analysis and associated visualizations were performed in R v4.1.3. Gene ontology (GO) enrichment analysis (Ashburner et al., 2000; https://geneontology.org): First, transcript identifiers were standardized by removing version suffixes to obtain ENSEMBL gene IDs. Genes with $\log_2$ fold change < 0 (downregulated) and those with $\log_2$ fold change > 0 (upregulated) were extracted separately. For each group, GO enrichment analysis was performed using the enrichGO function (https://rdrr.io/bioc/clusterProfiler/man/enrichGO.html) from the clusterProfiler package (v4.12.6; Wu et al., 2021), querying all three GO ontologies (biological process, cellular component, and molecular function) from the org.Hs.eg.db database (v3.19.1). Gene sets were filtered to include only those with a minimum of 3 and a maximum of 800 genes. A p-value cutoff of 0.05 was applied, with multiple testing correction performed using the false discovery rate (FDR) method. Gene set enrichment analysis (GSEA) (Subramanian et al., 2005; https://www.rdocumentation.org/packages/phenoTest/versions/1.20.0/topics/gsea) was

performed using the gseGO function (https://rdrr.io/bioc/clusterProfiler/man/gseGO.html), as implemented in the clusterProfiler package. Similarly, to GO, a ranked gene list was generated from differential expression results in a TSV file by standardizing gene identifiers (removing transcript version suffixes) and ordering the corresponding $\log_2$ fold-change values in descending order, yielding a vector of ENSEMBL gene IDs. This vector served as the input for the gseGO analysis, which was configured to consider all GO categories with 1,000 permutations to assess significance. Gene sets were filtered to include those with a minimum size of 3 and a maximum size of 800, and a p-value cutoff of 0.05 was applied without adjustment.

Visualization of functional enrichment results The enrichplot package (Yu G, 2025; enrichplot: Visualization of Functional Enrichment Result, R package version 1.26.6, https://yulab-smu.top/biomedical-knowledge-mining-book/) was employed to visualize functional enrichment results obtained from GSEA and GO analyses, including outputs from DOSE (Yu et al., 2015) and the clusterProfiler suite (Yu et al., 2012; Wu et al., 2021). Built entirely on the ggplot2 framework, enrichplot (https://yulab-smu.top/biomedical-knowledge-mining-book/enrichplot.html#heatmap-like-functional-classification) was used to generate publication-quality graphics that facilitated a comprehensive interpretation of the enrichment data. The results were visualized using custom R functions. The barplot function (https://www.rdocumentation.org/packages/graphics/versions/3.6.2/topics/barplot) was used to compute q-scores ($-\log_{10}$ of adjusted p-values) and generate a bar plot displaying the top enriched categories with customizable rounding and ggplot2 theming. The cnetplot function (https://rdrr.io/bioc/enrichplot/man/cnetplot.html) was applied to convert gene IDs from a TSV file into a ranked gene list, map them to gene names using the org.Hs.eg.db database (v3.19.1), and visualize the network of enriched gene sets with fold-change information. Lastly, the enrichment profile was visualized using traditional GSEA methods, the running score and preranked list, which were generated by the gseaplot function (https://rdrr.io/bioc/enrichplot/man/gseaplot.html) after specifying a gene set identifier, thereby summarizing the enrichment profile of the gene set of interest. Overall, **Tab. S13** provides a comprehensive

list of R packages, and their respective versions employed for functional analyses (GSEA and GO) and the visualization of enrichment results.

Table S13. List of R packages and associated versions utilized for functional enrichment analyses (GSEA and GO) and for visualizing the enrichment results as part of the RNA-seq data analysis and integration with eCLIP data

R package	Version
r-BiocManager	1.30.25
r-clusterProfiler	4.12.6
r-org.Hs.eg.db	3.19.1
r-tidyverse	2.0.0
r-magrittr	2.0.3
r-DOSE	3.30.5
r-enrichplot	1.24.4
r-scales	1.3.0
r-readr	2.1.5
r-stringr	1.5.1
r-dplyr	1.1.4
r-ggplot2	3.5.1
r-ggupset	0.4.1
r-AnnotationDbi	1.66.0

RESULTS

1. 12 h pre-me window was optimal for hPGCLCs specification

The model for hPGCLCs specification/early development used in this project was derived from the human embryonic stem cell male line WIS2 (Gafni et al., 2013) with a NANOS3-tdTomato knock-in reporter of PGC specification (**Fig. S1**) (Kobayashi et al., 2017). Notably, the *vitro*-generated PGC-like cells (PGCLCs) are transcriptionally similar to their *in vivo* PGC counterparts; hence, the nomenclature is used interchangeably to refer to PGCs prior to migration to the gonads.

The protocol based on the stepwise pre-me to PGC competence method by Kobayashi et al. (2017), described in Introduction section, was effectively applied in the laboratory (**Fig. 1**). Stem cells were induced to form pre-me cells, which exhibit transient competence for differentiation into PGCs, as evidenced by their acquisition of the expected morphology and formation of embryoid bodies (EBs) in 96-well U-bottom plates (**Fig. 1**). In the presence of cytokines (**Tab. S1**), the EBs express red fluorescence from NANOS3-tdTomato (**Fig. 1**), visualized by microscopy and sorted as detailed in **Fig. S10**, achieving a 20% differentiation efficiency (**Fig. 1)**. Optimization of pre-me incubation duration was conducted to maximize differentiation efficiency in-house with the cell line (**Tab. 1**). Similar to the findings by Kobayashi et al. (2017), unedited pre-me cells achieved optimal competence for differentiation into PGCs after a 12 h incubation period, yielding an average efficiency of 14.85%, which was superior to efficiencies observed at shorter or longer incubation times (**Tab. 1**). These results were obtained from three independent differentiation trials following the sorting protocol depicted in **Fig. S10**.

Table 1. Optimization of germ cell induction by evaluating different incubation durations during the pre-me stage. The data presented correspond to cells analysed and sorted on Day 4 of EB differentiation, based on the expression of the NANOS3-tdTomato marker. *The values represent differentiation efficiencies obtained from three independent experimental replicates.

Pre-me Incubation Length	Differentiation Efficiency of Day 4 sorted EBs*	Number of NANOS3 – tdTomato Positive Cells	Total Cell Count
3 h	1.66%	20,364	1,223,358
6 h	3.47%	78,846	2,275,295
12 h	**14.85%**	**196,123**	**1,318,599**
24 h	0.5%	2500	500,000
Negative Control (no cytokines)	0.02%	99	541,557

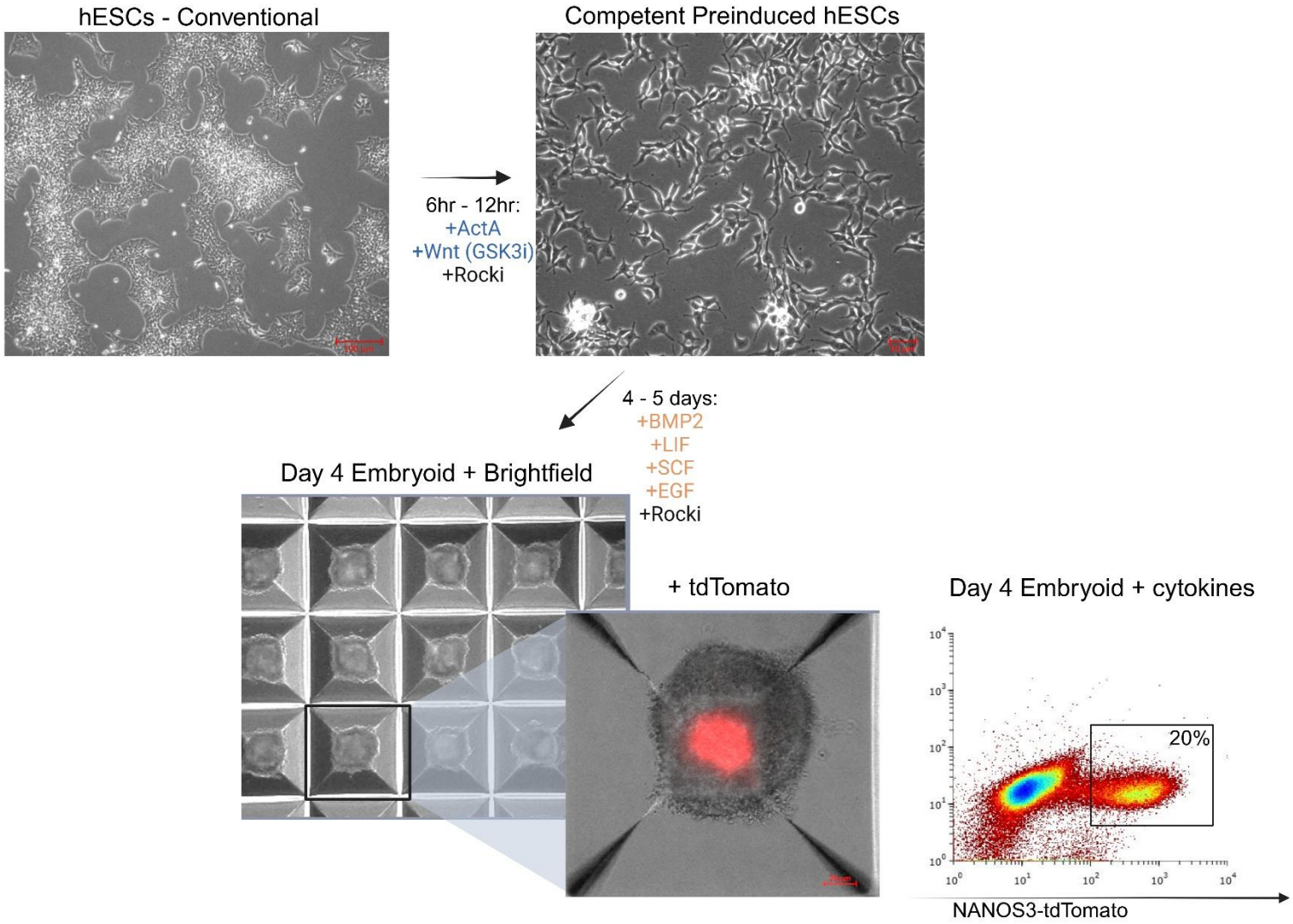


Figure 1. Workflow of the PGC specification protocol based on the stepwise pre-me to PGC transient competence method. Top left: Bright colonies of hESCs cultured in the specified culture medium. Scale bar, 100 µm. Top right: hESCs cultured with activin A and WNT activator to induce pre-me competence, displaying typical pre-induced morphology. Scale bar, 10 µm. Bottom: EBs cultured in 96-U bottom microwells (50x magnification), followed by a 200x magnified EB showing NANOS3-tdTomato signal. Scale bar, 20 µM. Bottom right: A sorting chart indicating the population of td-Tomato positive cells (20%). The fluorescence channel captures td-Tomato emission (581 nm) wavelength against side scatter area (SSC-A, logged) of sorted cells. In blue, pre-me cytokines: Activin A (Nodal Activator) and WNT (GSK3i, WNT pathway inhibitor). In yellow, PGCs cytokines: BMP2 (Bone morphogenetic protein 2); LIF (Leukemia inhibitory factor); SCF (Stem cell factor); EGF (Epidermal growth factor) and Rocki (Rho-associated protein kinase inhibitor (Kobayashi et al., 2017).

2. Endogenous NANOS1 mRNA and protein are expressed in the pre-me stage of differentiation of the W15 stem cell line toward PGCs

First, endogenous NANOS1 mRNA levels were assessed by RT-qPCR targeting the 3′UTR across various stages of W15 stem cell line differentiation, using the stem cell stage (Day -1) as a reference (**Fig. 2A**). The highest NANOS1 mRNA levels were detected at Day 0 (pre-me) and on Day 4 of EB differentiation, in

particular, within the germ cell population (Day 4 Germ, PGCs), while the sorted somatic cell population exhibited the lowest expression levels (**Fig. 2A**).

In the absence of a specific antibody for human NANOS1, a 3x-FLAG tag was inserted at the C-terminus of NANOS1, just upstream of the STOP codon in its single exon, using CRISPR-Cas9 technology. The detailed list of knock-in components, primers, and guide RNAs is provided in **Tab. S2** and **S3**, as well as **Fig. S6**. Following CRISPR editing and subsequent cell sorting (as detailed in the Methods section and illustrated in **Fig. S9**), among the nine generated cell lines only the NANOS1- L3 cell line uniquely demonstrated homozygosity for the knock-in, as determined by PCR using primers flanking the edited region (**Tab. S3**; **Fig. 2B**). The higher molecular weight band observed in the NANOS1-L3 cell line compared to the unedited W15 control (293 bp with FLAG versus 223 bp) confirmed successful gene editing, and this was further validated by Sanger sequencing.

This expression pattern was corroborated by RT-qPCR targeting the 3×-FLAG sequence in the CRISPR-edited NANOS1-L3 cell line, which showed a significant upregulation at the pre-me stage compared to the stem cell stage (**Fig. 2C**). Additionally, NANOS1 RNA levels in unsorted Day 4 EBs were comparable to those observed at the stem cell stage, underscoring the opposing trends in soma and germ cell populations (**Fig. 2A and Fig. 2C**).

Beyond genotyping, the NANOS1-L3 cell line was evaluated for its differentiation potential toward PGCLCs using flow cytometry. Notably, the NANOS1-L3 cell line achieved optimal differentiation efficiency with a shorter, 3 h pre-me incubation (**Fig. 2D**), compared to the 12 h pre-me required by the unedited W15 line (**Tab. 1**). In the presence of cytokines, the NANOS1-L3 cell line reached a differentiation efficiency of 27.6% by Day 4 relative to the negative control (without cytokines) (**Fig. 2D**), confirming its suitability for further protein characterization.

Given that the pre-me stage exhibits the highest levels of endogenous NANOS1 mRNA expression (**Fig. 2A and Fig. 2C**) and allows for the collection of sufficient cell numbers, unlike the PGC stage (Day 4), endogenous NANOS1 protein levels were evaluated by immunoprecipitation from pre-me stage followed by western blotting using an antibody targeting the 3x-FLAG tag (**Fig. 3A-D**). The immunoprecipitation results (**Fig. 3A-B**) confirmed the presence of endogenous NANOS1 protein during the pre-me stage of W15 ell line differentiation. The specificity of the FLAG antibody was also demonstrated, as it did not bind to untagged NANOS1 in the W15 unedited negative control. Instead, it showed a signal for the NANOS1-3xFLAG protein in a W15 cell line (WT 1.2.5.1-N13xFLAG-TetON), described below, where NANOS1 is inducibly overexpressed via the TET-ON system incorporated within a PiggyBac transposon (**Fig. 3A-C**). Notably, comparison of endogenous NANOS1 levels to those of the inducibly overexpressed counterpart reveals that the endogenous protein is present at relatively low abundance in this model system particularly at the pre-me stage (**Fig. 3A-C**). These findings indicate that, although endogenous NANOS1 protein is expressed in the model system, its levels are relatively low, potentially limiting the ability to observe the functional roles of the endogenous protein and its MUT form in PGCLCs.

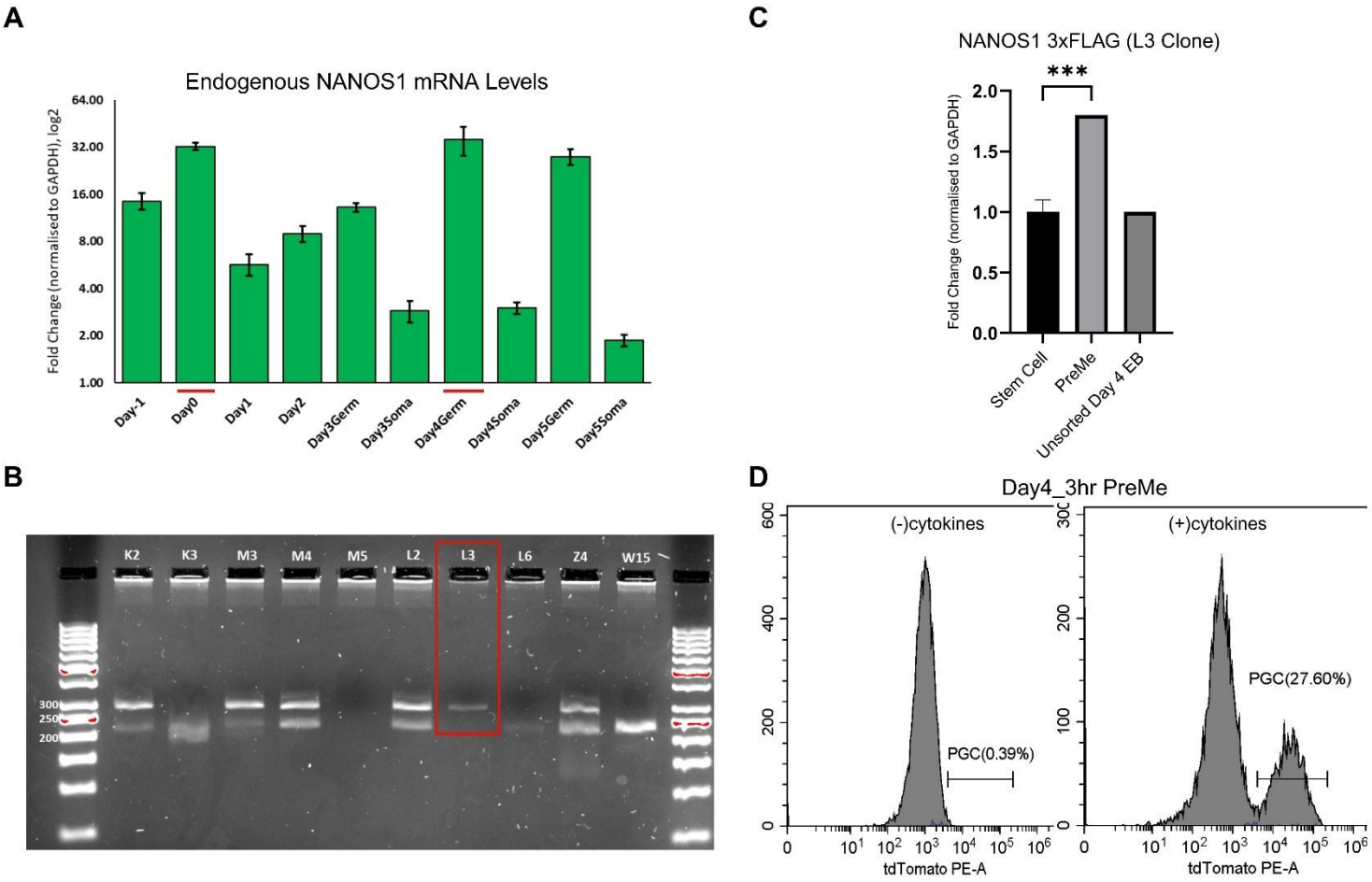


Figure 2. Highest endogenous NANOS1 mRNA levels in pre-me and PGCs before and after CRISPR-Cas9 knock-in of the 3×-FLAG in the NANOS1 sequence. A- Genotyping of CRISPR edited clones: 200 ng of DNA from clones K2–L2 and L3–Z4 were resolved on a 2% agarose gel. Additionally, 86 ng of DNA from L3 and 137 ng from L6 were loaded. The W15 clone served as an unedited negative control. **B**- RT-qPCR analysis of the W15 stem cell line at various differentiation stages: Day –1 represents the stem cell stage; Day 0 represents pre-me; and Days 1 to 5 correspond to embryoid body (EB) differentiation, further subdivided into germ and soma populations. Expression levels (log2) were normalized to GAPDH, with experiments performed in triplicate. Primers targeted the 3′UTR of NANOS1 mRNA (Forward: TCCGTGCTGAACGATTGGGA; Reverse: CTTCGCCGTTCTGGATGTGC). **C**- RT-qPCR analysis of the CRISPR edited NANOS1 L3 clone at the stem cell, pre-me, and unsorted Day 4 EB stages. Expression levels were normalized to GAPDH, with experiments performed in triplicate. Statistical analysis was conducted using one-way ANOVA, with the stem cell stage as the reference. Primers targeted the 3×-FLAG knocked-in sequence (Forward: GGCAAGAAGCTGCGCGACTA; Reverse: CGGGCCTTCACTTGTCATCG). **D**- Flow cytometric histogram analysis of the L3 clone after EB digestion on Day 4 of differentiation following a 3 h pre-me incubation. The left histogram represents the EB negative control without cytokines, while the right histogram shows the EB treated with cytokines, highlighting the td-Tomato–positive germ cell population at 27.60%.

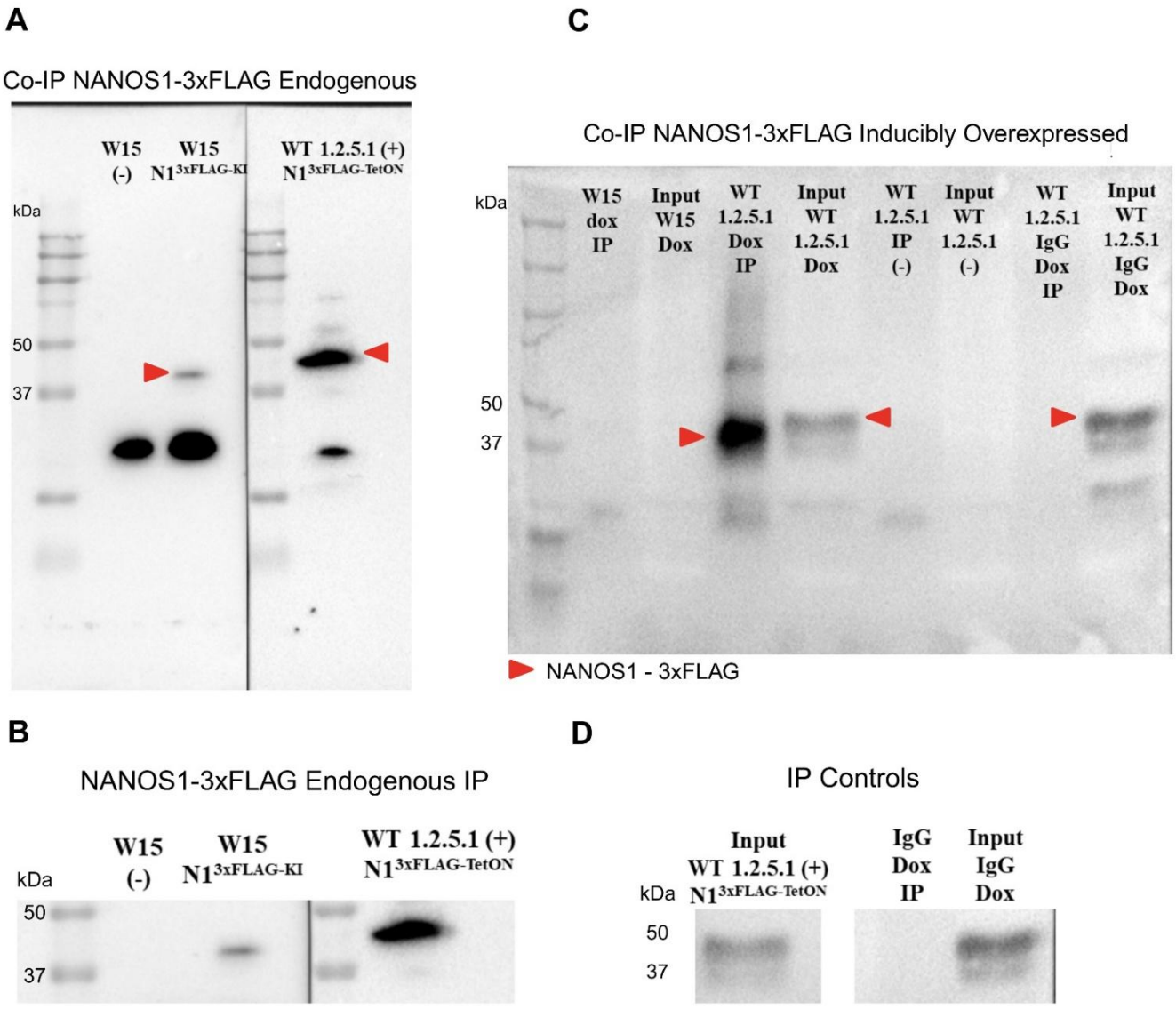


Figure 3. Co-Immunoprecipitation of NANOS1-3xFLAG in CRISPR-edited cell line L3 at the pre-me stage. A- western blot analysis of co-immunoprecipitation (co-IP) at the pre-me stage was performed on three cell lines: the negative control W15 (12 mln cells), the L3 cell line (W15 1.2.5.1-N13xFLAG-KI, 18 mln cells), and the positive control PiggyBac cell line (WT 1.2.5.1-N13xFLAG-TetON, 12 mln cells). The L3 cell line was exposed for 300 seconds using an anti-FLAG antibody at a 1:1250 dilution, whereas the positive control was exposed for 3 seconds using an anti-FLAG antibody at a 1:1500 dilution. **B**- The highlighted bands representing the NANOS1-3xFLAG protein from the corresponding upper blot. **C**- western blot analysis of co-IP for negative and positive control cell lines, both with and without doxycycline (dox) to induce NANOS1-3xFLAG overexpression. To confirm the specificity of the FLAG antibody, an IgG antibody was used in equivalent amounts to the FLAG antibody during IP. **D**- A zoomed-in view of the input for the positive control cell line, as well as the IP and input for the IgG control of the same cell line. A 4-15% resolving gel was utilized to visualize the protein of interest.

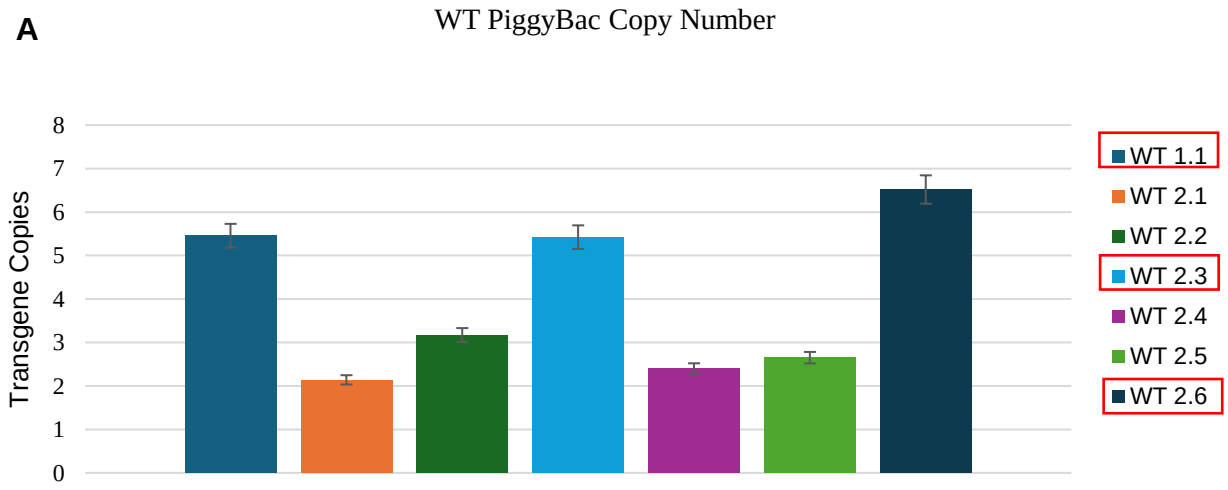
3. PGCs differentiation from established WT- and MUT-NANOS1 edited hESC cell lines

Given the previously observed low endogenous levels of NANOS1, edited cell lines were engineered to overexpress both WT- and MUT-NANOS1 p. [Pro34Thr; Ser78del] variant. The PiggyBac and Tet-ON systems were utilized together to induce NANOS1 protein overexpression, facilitating selection,

visualization, and immuno-capture assays. Incorporating TET-ON plasmids for MUT-NANOS1 requires a hypbase plasmid to transpose the Tet-On sequence into the genome (**Fig. S3**). This inducible system is preferred over noninducible overexpression as stem cells undergo epigenetic resetting during germ cell pathway commitment; too early overexpression might inhibit differentiation.

3.1 PGCs differentiation challenge from established WT- and MUT-NANOS1-edited hESC cell lines via antibiotic selection and FACS

The design of transposon plasmids for the initial approach is shown in **Fig. S4A-B**, and the initial cloning protocol is detailed in the Methods section and **Fig. S15A-C**. Due to the nature of PiggyBac editing, transgene cassette copy numbers were determined using established protocols (Lin et al., 2014). The protocol and standard curves for determining transposon absolute copy numbers are detailed in **Tab. S8** and **Fig. S16A-B**. As shown in **Fig.4A-B**, WT-NANOS1 cell line WT 1.1 (5.5 copies), WT 2.3 (5.5 copies), and WT 2.6 (6.5 copies) (**Fig. 4A**), along with NANOS1 MUT cell lines MUT 1.1 (7.5 copies), MUT 2.2 (8.5 copies), MUT 2.7 (9 copies), and MUT 2.8 (approx. 12 copies) (**Fig. 4B**), had the highest copy numbers of inserted Tet-On cassettes. These NANOS1 cell lines are expected to exhibit higher transgene protein expression levels when stimulated with dox.



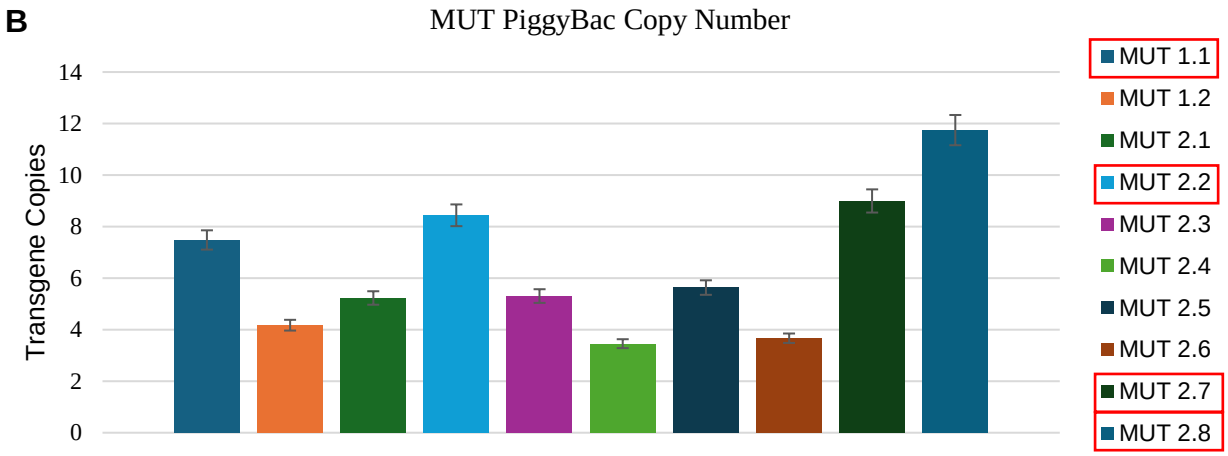


Figure 3. NANOS1 absolute transgene copy numbers in WT- and MUT-NANOS1 cell lines. A- Transgene copy numbers for 7 WT-NANOS1 cell lines determined using the standard curve (**Fig. 1S6A**). **B**- Transgene copy numbers for 10 MUT-NANOS1 cell lines determined using the standard curve (**Fig. 16B**). Cell lines with the highest copy numbers are in red squares.

Additionally, the high copy number cell lines from **Fig.4A-B** were further evaluated to determine NANOS1-3xFLAG inducible protein overexpression using western blot. The overexpressed protein, detected with a 3x-FLAG antibody, was successfully detected at comparable levels across all tested cell lines (**Fig. 5**). The protein expression results mirrored the copy number analysis (**Fig. 4A-B** and **Fig. 5**), with MUT-NANOS1 cell lines showing higher levels of NANOS1-3xFLAG, corresponding to transgene cassette copies ranging from 7.5 to 12. Importantly, no protein expression was detected without dox induction, indicating tight system control (no "leakage"). Moreover, the antibody successfully detected the fused NANOS1(3xFLAG)-P2A-mVenus protein, demonstrating that the P2A cleavage peptide did not achieve complete separation between tagged NANOS1 and the fluorescence marker mVenus.

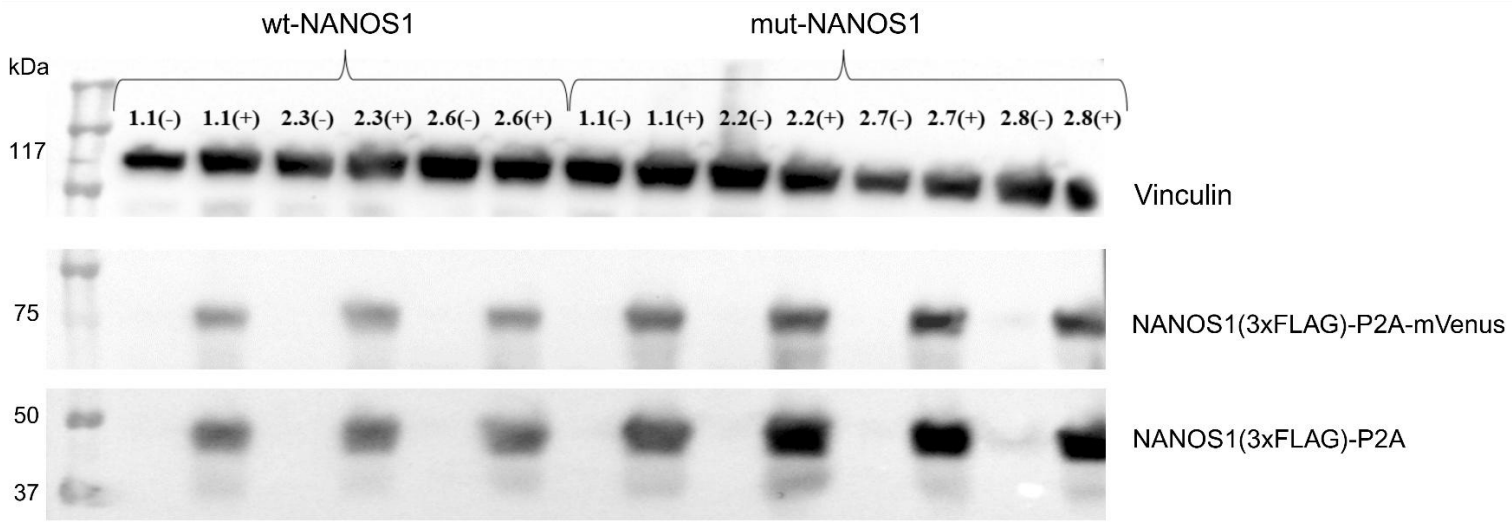


Figure 4. Validation of inducible NANOS1 protein overexpression in WT- and MUT-NANOS1 cell lines. For each cell line, samples treated with dox were compared with their corresponding untreated negative controls. Western blot analysis was performed using anti-FLAG antibody (1:1500) to detect the NANOS1 fusion proteins. The middle panel shows the fused NANOS1-3xFLAG-mVenus protein (~75 kDa) detected at a 57.3'' exposure, while the lower panel (bottom image) displays the NANOS1-3xFLAG-P2A tagged protein at a 3'' exposure. A reference for the housekeeping protein, vinculin (117kDa), detected at a 1:10000 dilution with a 3'' exposure, is shown at the top of the blot.

The mVenus fluorescence marker, incorporated into the transgene cassette via the P2A linker, was used to evaluate the editing efficiency of the protocol using puromycin selection with cloning cylinders (protocol in **Fig. S15A-C**).

Overall, MUT-NANOS1 cell lines exhibited superior transgene cassette translational efficiency, with three out of four cell lines surpassing 90% efficiency (as assessed by mVenus fluorescence). Among these, MUT 2.7 demonstrated the highest efficiency at 94.38% (**Fig. 6**). In contrast, WT-NANOS1 cell lines showed lower editing efficiency, as indicated by the histograms in **Fig. 6**, where the FITC channel was used to capture the mVenus fluorescence marker's emission wavelength at 527 nm. Among the high copy number WT-NANOS1 cell lines, WT 1.1 demonstrated the highest editing efficiency, with 94.86% of cells exhibiting mVenus fluorescence (**Fig. 6**). Notably, the improved editing efficiencies observed in the cell lines highlighted in **Fig. 6** were achieved after a second round of puromycin selection combined with FACS enrichment, resulting in elevated NANOS1-mVenus expression in some cell lines. Initially, a major limitation of the first round of puromycin selection was its low editing efficiency, as indicated by weak mVenus expression (data not shown) across the cell lines. However, additional puromycin (antibiotic) treatment and selection may have imposed cellular stress, potentially compromising stemness and pluripotency, thus potentially diminishing the cell lines' ability to differentiate in response to cytokines. Subsequently, the cell lines failed to differentiate at 12 h pre-me (data not shown), which was the optimal timing for the unedited W15 stem cells (**Tab. 1**). As a result, they were subjected to earlier pre-me induction times, under the assumption that cellular stress may have caused them to lose pluripotency sooner than their unedited counterparts.

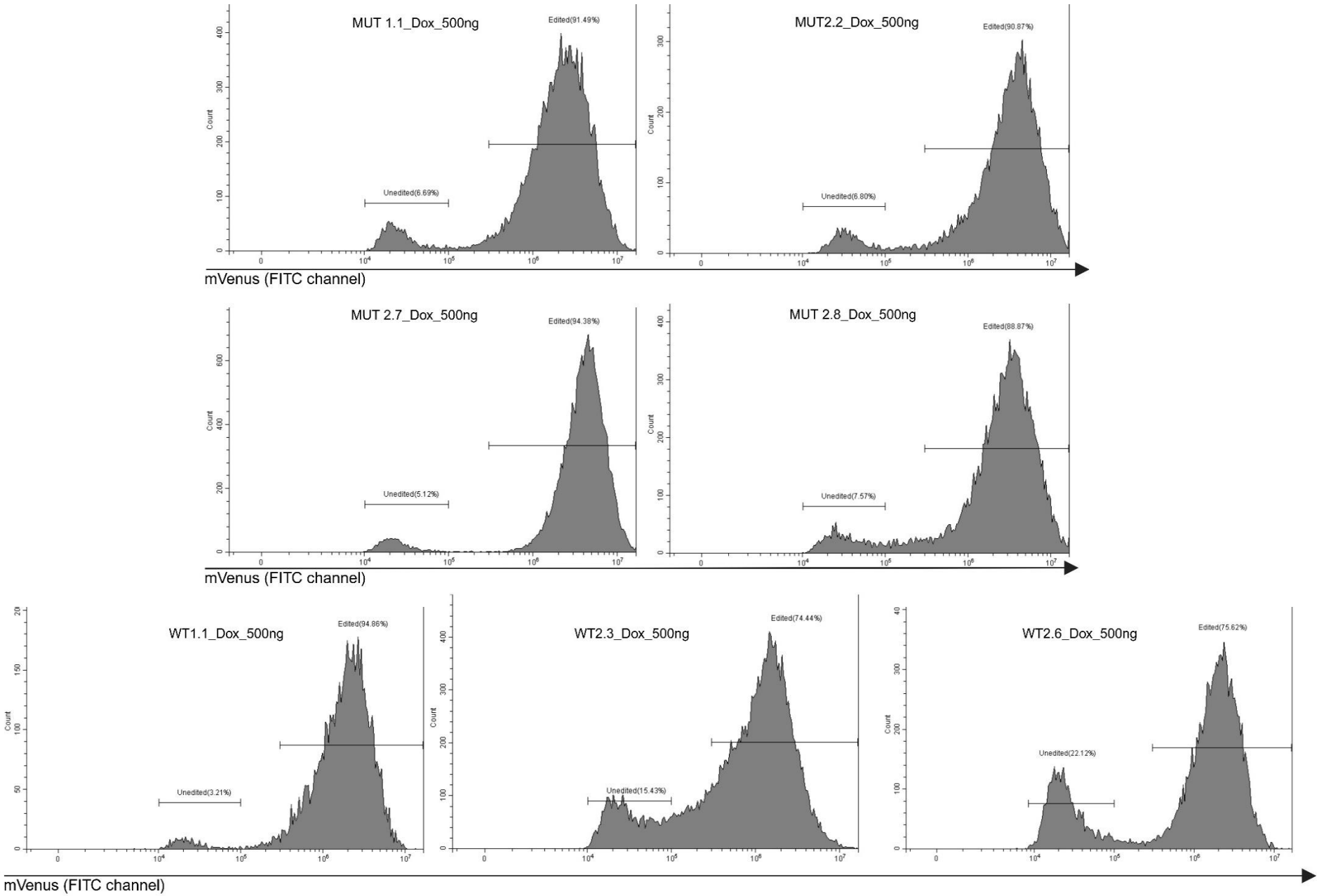

MUT 1.1_Dox_500ng
Edited(91.49%)
Unedited(6.69%)
MUT2.2_Dox_500ng
Edited(90.87%)
Unedited(6.80%)
Count
mVenus (FITC channel)
MUT 2.7_Dox_500ng
Edited(94.38%)
Unedited(5.12%)
MUT 2.8_Dox_500ng
Edited(88.87%)
Unedited(7.57%)
Count
mVenus (FITC channel)
WT1.1_Dox_500ng
Edited(94.86%)
Unedited(3.21%)
WT2.3_Dox_500ng
Edited(74.44%)
Unedited(15.43%)
WT2.6_Dox_500ng
Edited(75.62%)
Unedited(22.12%)
Count
mVenus (FITC channel)

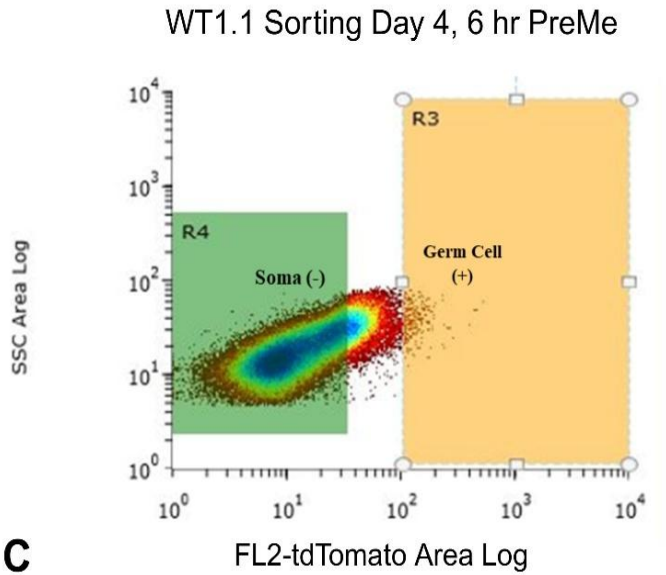
Figure 6. Flow cytometry analysis of the mVenus marker within the NANOS1 transgene. The histograms display the unedited (negative) cell population on the left and the positive (edited) cell population on the right, expressing mVenus as part of the WT- and MUT- NANOS1-3xFLAG-P2A-mVenus cassette. The marker's fluorescence is detected using the FITC channel (emission wavelength 528 nm). The histograms represent clones that were re-selected with puromycin for a second round and enriched via FACS sorting targeting mVenus fluorescence following transgene copy number analysis (**Fig. 4**).

To determine the optimal pre-me window for successful PGC differentiation, pluripotency markers (KLF4, NANOG, OCT4, SOX2, and MYC) and primitive streak/pre-me markers (Brachyury - T, MIXL, and EOMES) were quantified using RT-qPCR. The stem cell stage (Day-1) served as the reference point for these comparisons following induction into the pre-me stage. The expected fold-changes were compared to those reported by Kobayashi et al. (2017), who quantified germ cell induction from pre-me at 6 h intervals and identified peak competency at 12 h, a trend similar to that observed in the unedited W15 cell line differentiation trials (**Tab. 1**). According to the literature, germ cell-competent pre-me cells typically maintain equal or higher levels of pluripotency markers relative to the stem cell stage, except for SOX2, which undergoes a slight reduction (Kobayashi et al., 2017). In contrast, only moderate, not pronounced, upregulation of primitive streak and pre-me markers was observed, as higher levels of these markers would indicate mesendoderm (ME) induction, occurring after the window of peak germ cell induction competency (Kobayashi et al., 2017).

WT-NANOS1 cell line WT 1.1, which exhibited the highest editing efficiency based on mVenus expression levels (**Fig. 6**), was subjected to differentiation efficiency testing and RT-qPCR analysis. As shown in **Fig. 7B**, the clonal pre-me cells were tested at both 3 h and 6 h induction windows, after which cytokines for germ cell differentiation were added. At 6 h pre-me, the pluripotency markers KLF4, OCT4, and MYC were higher than in stem cells, as expected, while on the contrary NANOG expression decreased, albeit not significantly (**Fig. 7B**). SOX2 and primitive streak markers followed expected trends, with SOX2 moderately decreasing and primitive streak markers moderately increasing at 6 h pre-me compared to 3 h (**Fig. 7B**).

The WT 1.1 cell line was sorted using the td-Tomato marker to assess PGC differentiation efficiency (FL2 channel, S3e™ Cell Sorter, Bio-Rad) at 6 h pre-me based on RT-qPCR data (**Fig. 7A-B**). As highlighted in both **Fig. 7A** and **Fig. 7C**, the differentiation efficiency was low, averaging 0.20%, with an average of only 306 positive sorted cells per differentiation. These values reflect the low levels of NANOG expression (**Fig. 7B**), which, along with OCT4, are typically found at high levels in T posterior-stratified epiblast cells that are competent to give rise to a germ cell subpopulation. NANOG expression remains elevated when this population emerges (Kobayashi et al., 2017).

A



C

Label	Count	Mean	% of Total
Total	**137493**	16.80	91.98
R4	122119	12.77	81.69
R3	**306**	136.23	**0.20**

B

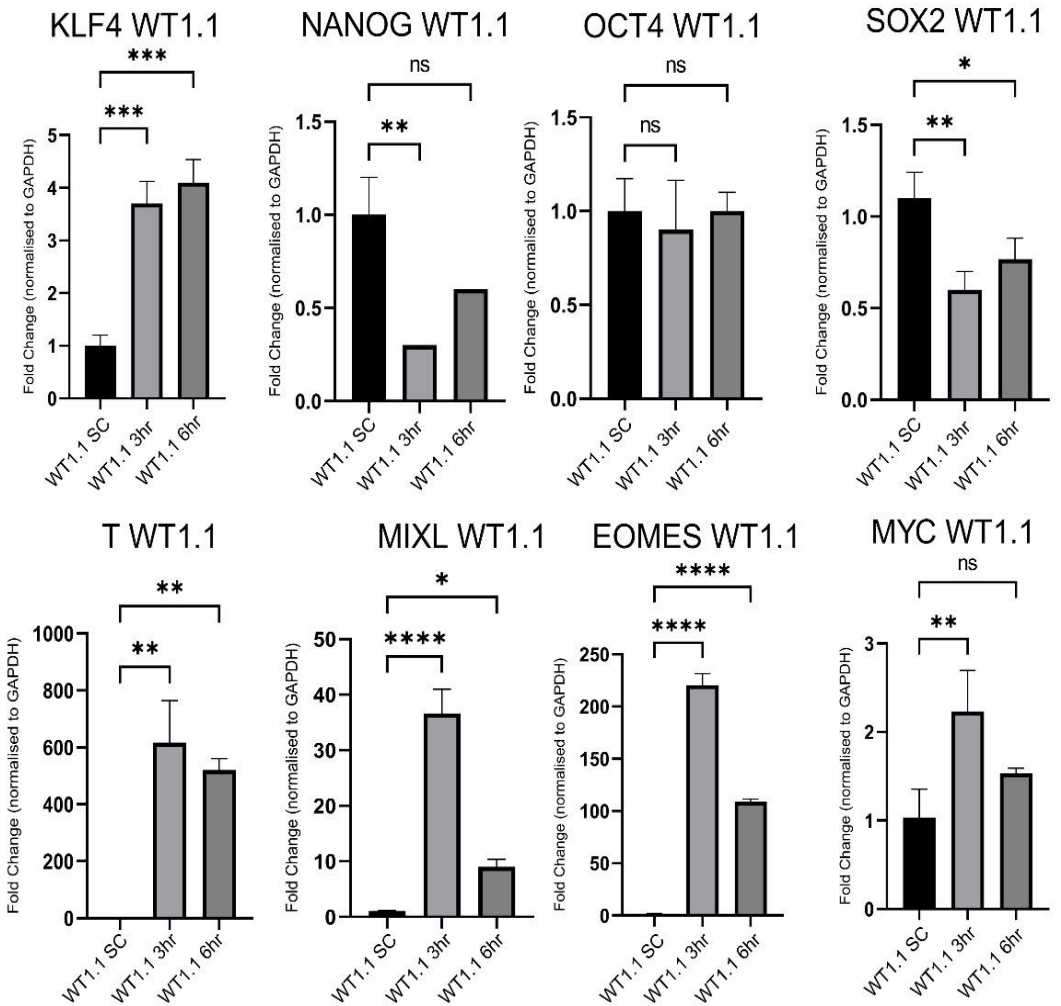


Figure 5. Evaluating the differentiation efficiency of WT-NANOS1 cell line 1.1 into PGCLCs. A- Schematic of fluorescence-based cell sorting using the tdTomato marker fused to NANOS3. R4 indicates the negative somatic cell population, while R3 gates the positive tdTomato-expressing germ cell population. The cells were differentiated at 6 h post pre-me competence induction. **B**- RT-qPCR analysis of pluripotency and primitive streak markers, based on the previous study by Kobayashi et al. (2017). Values were derived from 3 biological replicates, and statistical significance was determined using an unpaired Student's t-test with p-values < 0.05 considered significant. SC represents stem cells (Day -1); 3 h indicates 3 h pre-me induction; 6 h indicates 6 h pre-me induction. Fold changes were normalized against GAPDH. **C**- Sorting table following FACS, as detailed in A, from 3 separate differentiation trials. It includes cell counts based on specific gating, mean values, and the percentage of each gated population relative to the total sorted cells.

The MUT-NANOS1 cell line 2.7, which exhibited the highest editing efficiency among the generated MUT cell lines based on mVenus expression levels (**Fig. 6**), was initially subjected to RT-qPCR analysis for pluripotency markers (KLF4, NANOG, OCT4, SOX2, and MYC) and primitive streak/pre-me markers (T, MIXL, and EOMES) at both 3 h and 6 h pre-me induction time points. As shown in **Fig. 8**, except for KLF4 and MYC, the levels of pluripotency markers were lower than those in stem cells, with SOX2 being significantly reduced. Additionally, the levels of T, MIXL, and EOMES were much higher than anticipated (**Fig. 8**), suggesting that the edited stem cells were not sufficiently competent and pluripotent to differentiate at the pre-me stage. Consequently, two other MUT-NANOS1 cell lines with high editing efficiencies, specifically cell lines 1.1 and 2.2 (**Fig. 6**), were further tested for their differentiation efficiency toward PGCLCs (**Fig. 9** and **10**).

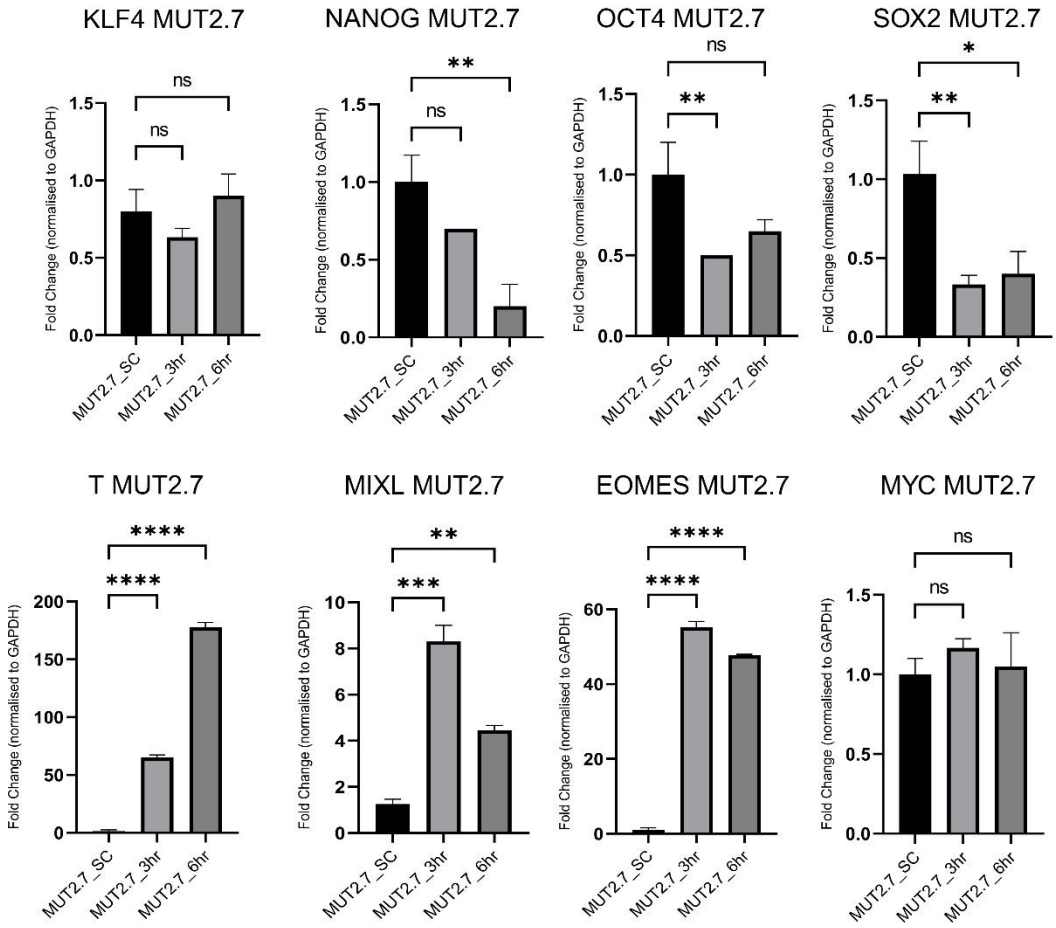
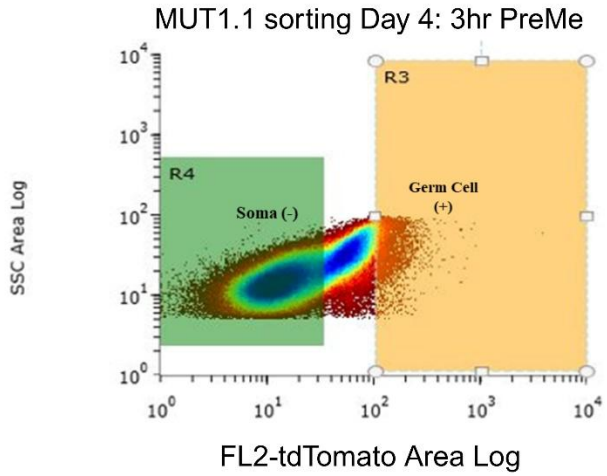


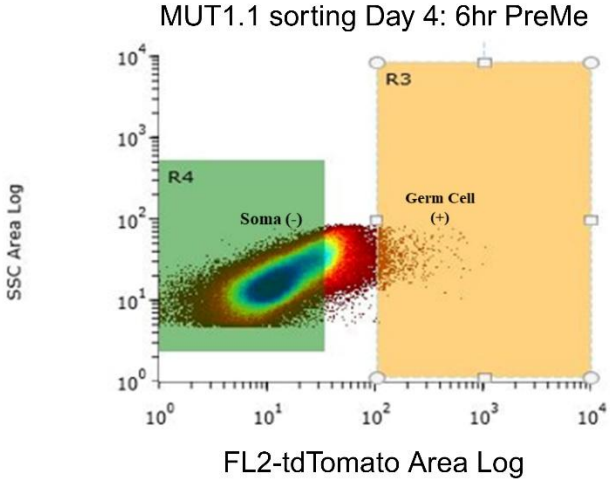
Figure 8. RT-qPCR analysis of pluripotency and primitive streak markers expressed by MUT-NANOS1 2.7. The panel of markers is based on previous analyses by Kobayashi et al. (2017). Values were derived from 3 biological replicates. Statistical significance was determined using an unpaired Student's t-test with p-values < 0.05 considered significant. SC refers to stem cells (Day -1); 3 h indicates 3 h of pre-me induction; 6 h indicates 6 h of pre-me induction. Fold changes were normalized to *GAPDH*.

In the case of MUT-NANOS1 cell line 1.1, the levels of pluripotency factors remained high at both 3 h and 6 h pre-me, except for SOX2, which showed a more significant decrease at 6 h compared to 3 h (**Fig. 9B**). Similarly, the primitive streak/pre-me markers (T and EOMES) were relatively high compared to the stem cell counterpart, except for MIXL, which sharply decreased at 6 h (**Fig. 9B**). The cell line underwent differentiation trials at both pre-me times, shown in **Fig. 9A** (3 h window) and **Fig. 9C** (6 h window). On average, the sorting efficiency was moderately better at 3 h pre-me, with a differentiation efficiency of 1.17% and 6 602 tdTomato-positive sorted cells (**Fig. 9D**). A later pre-me stage of 6 h did not improve overall differentiation at 0.15% or increase the number of sorted cells (average 811 cells). Despite the promising pluripotency marker profile (**Fig. 9B**) and high levels of mVenus expressing cells (**Fig. 6**), MUT-NANOS1 1.1 cell line had higher than usual levels T and EOMES, indicating a tendency toward ME lineages. This resulted in low differentiation efficiency at the earliest 3 h pre-me, with decreasing trends toward 6 h and no differentiation at 12 h (data not shown).

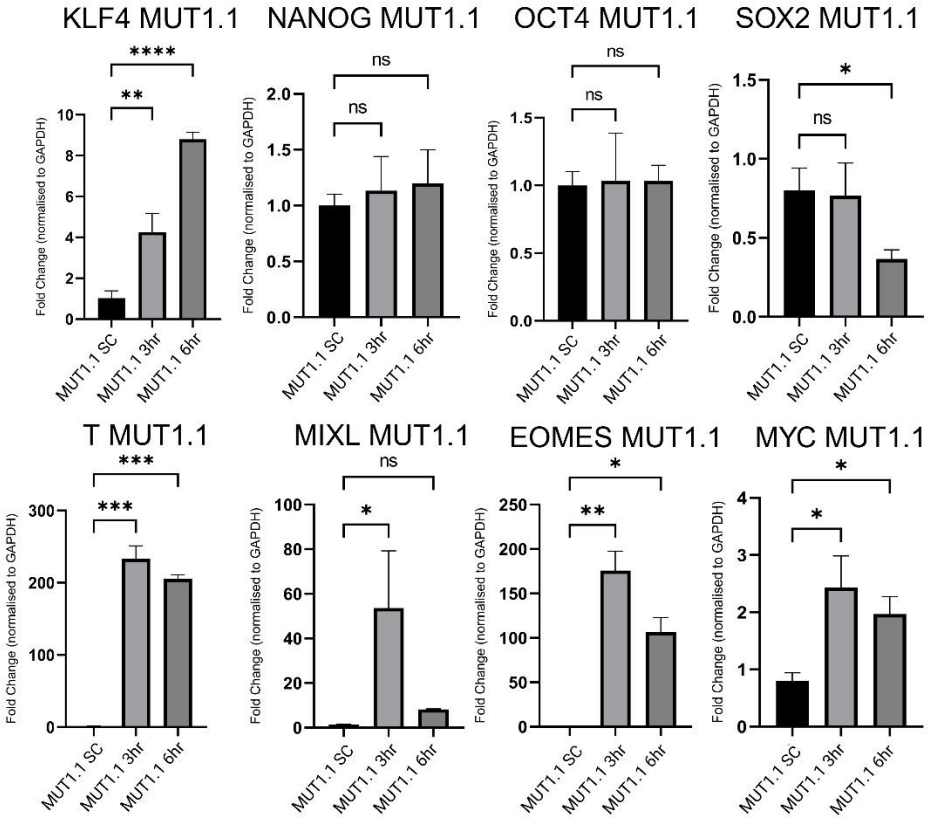
A



C



B



D

3hr PreMe

Label	Count	Mean	% of Total
Total	**522093**	27.73	92.43
R4	374281	14.49	66.26
R3	**6602**	135.50	**1.17**

6hr PreMe

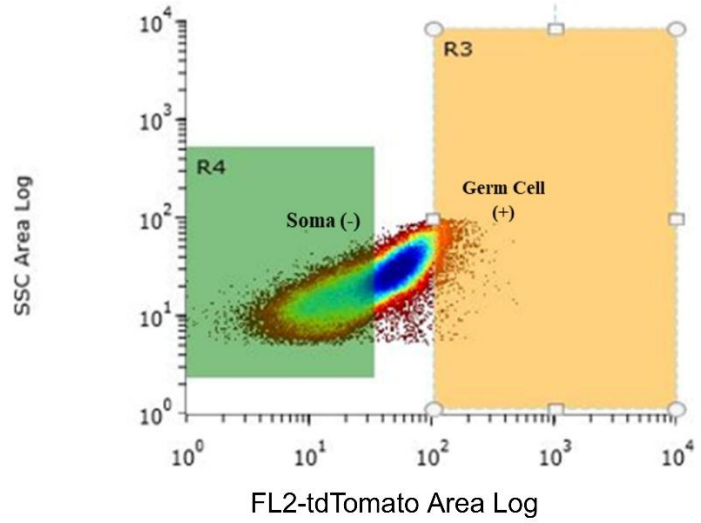
Label	Count	Mean	% of Total
Total	**501935**	15.55	91.09
R4	471356	13.39	85,54
R3	**811**	162.87	**0.15**

Figure 9. Assessing differentiation efficiency of MUT-NANOS1 cell line 1.1 into PGCs. A- Diagram of fluorescence-based cell sorting using the tdTomato marker fused to NANOS3. R4 represents the negative somatic cell population, while R3 gates the positive tdTomato-expressing germ cell population. Cells were sorted on Day 4 of EB differentiation following 3 h of pre-me induction. **B**- RT-qPCR analysis of pluripotency and primitive streak markers based on the panel determined by Kobayashi et al. (2017). Values were obtained from 3 biological replicates. Statistical significance was determined using an unpaired Student's t-test with p-values < 0.05 considered significant. SC refers to stem cells (Day -1); 3 h indicates 3 h of pre-me induction; 6 h indicates 6 h of pre-me induction. Fold changes were normalized to *GAPDH*. **C**- Diagram of fluorescence-based cell sorting using the tdTomato marker fused to NANOS3, similar to part A. Cells were sorted on Day 4 of EB differentiation following 6 h of pre-me induction. **D**- Sorting tables from FACS analysis as detailed in A and C from 3 separate differentiation trials after 3 h (top) and 6 h (bottom) of pre-me induction. The tables show cell counts based on specific gating, mean values, and the percentage of each gated population relative to the total population sorted.

Lastly, for MUT-NANOS1 cell line 2.2, only the 3 h pre-me window was tested (**Fig. 10A-C**) due to the overall lower levels of pluripotency markers (KLF4, NANOG, OCT4, SOX2 and MYC) observed at 6 h pre-me (**Fig. 10B**). Conversely, the primitive streak markers and EOMES showed higher than expected levels within the 3 h timeframe compared to the 6 h counterpart (**Fig. 10B**). Following sorting as depicted in **Fig. 10A**, the average differentiation efficiency for this clone was 1.5%, yielding 2 285 tdTomato-positive germ cells sorted on average (**Fig. 10C**). The efficiency correlated with the higher levels of pluripotency observed via RT-qPCR (**Fig. 10B**). However, similar to MUT-NANOS1 cell line 1.1, the levels of primitive streak/pre-me markers (T, MIXL and EOMES) were significantly higher than published (Kobayashi et al., 2017), potentially altering the cell response to BMP2/4 signaling, which may push the cells towards a mesodermal fate and reduce their propensity to enter the PGC differentiation pathway.

Overall, the first set of PiggyBac edited cell lines did not achieve the required differentiation efficiencies for downstream cell sorting and collection for further analysis. This failure was attributable to limitations in editing efficiency, as evidenced by the low percentage of cells expressing the introduced transgene cassette. Consequently, antibiotic reselection and FACS enrichment were necessary, which subsequently compromised the pluripotency and stemness of the edited cell lines. Additionally, differences in pre-me timing compared to the control cell line W15, along with the absence of the expected pluripotency and primitive streak marker profiles (evaluated by KLF4, NANOG, OCT4, SOX2, T, MIXL, EOMES, and MYC), prevented the competent pre-me cells from differentiating into the PGCLC lineage in response to BMP2 and other cytokines.

A



C

Label	Count	Mean	% of Total
Total	**142640**	44.60	93.55
R4	48842	18.79	32.03
R3	**2285**	126.25	**1.5**

B

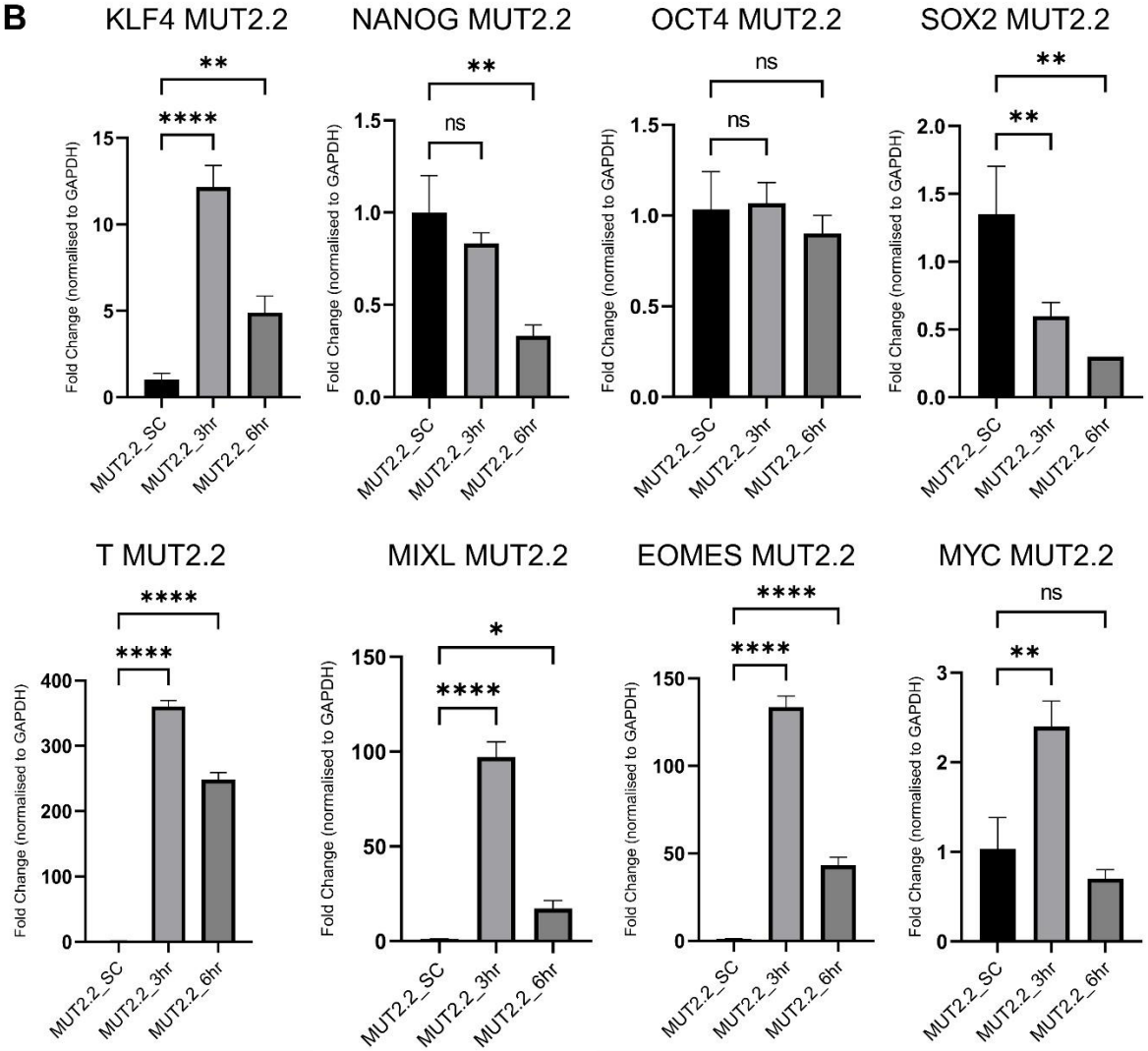


Figure 6. Assessing differentiation efficiency of MUT-NANOS1 cell line 2.2 into PGCs. A- Diagram of fluorescence-based cell sorting using the tdTomato marker fused to NANOS3. R4 represents the negative somatic cell population, while R3 gates the positive tdTomato-expressing germ cell population, captured via the FL2 channel against the side scatter area log (SSC Area Log). Cells were sorted on Day 4 of EB differentiation following 3 h of pre-me induction. **B**- RT-qPCR analysis of pluripotency and primitive streak markers, based on the panel determined by Kobayashi et al. (2017). Values were obtained from 3 biological replicates. Statistical significance was determined using an unpaired Student's t-test with p-values < 0.05 considered significant. SC refers to stem cells (Day -1); 3 h indicates 3 h of pre-me induction; 6 h indicates 6 h of pre-me induction. Fold changes were normalized to *GAPDH*. **C**- Sorting table from FACS analysis, as detailed in A, from 3 separate differentiation trials following 3 h of pre-me induction. The table shows cell counts based on specific gating, mean values, and the percentage of each gated population relative to the total population sorted.

3.2 Four stable (2 WT- and 2 MUT-NANOS1) cell lines generated via 96-well cell seeding and FACS selection able to differentiate into PGCs

The second approach to edited cell line generation utilized fluorescence-activated cell sorting (FACS) followed by cell seeding in 96-well plates instead of antibiotic selection. This method was carried out in the presence of Rocki (ROCK inhibitor) in the media to minimize cell apoptosis, at low sorting speeds, and by maintaining the plates at a stable 37°C temp. in the incubator before and after sorting. These conditions enabled the generation of stable, pluripotent cell lines for further analysis. Combined with the automated precision of cell sorting in 96-well plates, this approach successfully generated edited cell lines, overcoming the inefficiencies of the previous method. The prior approach relied on antibiotic (puromycin) selection incorporated in the cassette, the use of cloning cylinders that may not have captured homogeneous stem cell colonies, and repeated FACS enrichment. These factors together induced cellular stress, leading to decreased pluripotency and stemness, as evidenced by the low levels of PGC differentiation among the tested cell lines (**Figs. 7–10**). Furthermore, the second approach used a simplified PiggyBac transposon design that incorporated superfold-GFP (sfGFP) instead of mVenus in the NANOS1 cassette and omitted the antibiotic resistance cassette, as puromycin selection was not applied (see **Methods**: **Figs. S5** and **S7**). Flow cytometry analysis of the resulting cell lines (**Fig. 11**) demonstrated a robust and uniform response to dox induction, with over 90% of the cells expressing the sfGFP-NANOS1 signal. Detailed measurements, including the number of cells seeded per cell line, are provided in **Tab. 2**. Due to the low efficiency of single-cell seeding in stem cell cloning, three cells were seeded in duplicate wells using the same transposon-to-transposase ratio. As a result, most of the cell lines selected for further analysis were derived from the 3-cell seeding protocol. Initially, four WT-NANOS1 cell lines were chosen based on sfGFP intensity (1.1.1, 3.1.1, 1.2.5.1, 3.10.1) (**Fig. 11**), with sfGFP-expressing stem cell percentages post-dox induction ranging from 95% to 98% (**Tab. 2**). A greater number of MUT-NANOS1 cell lines were initially selected (1.1.2, 3.1.5, 3.1.8, 3.2.5.2, 3.5.1, 3.5.3) (**Fig. 11**), showing sfGFP intensity comparable to the WT-NANOS1 cell lines, although their stem cell percentages were slightly lower, ranging from 90% to 94% (**Tab. 2**).

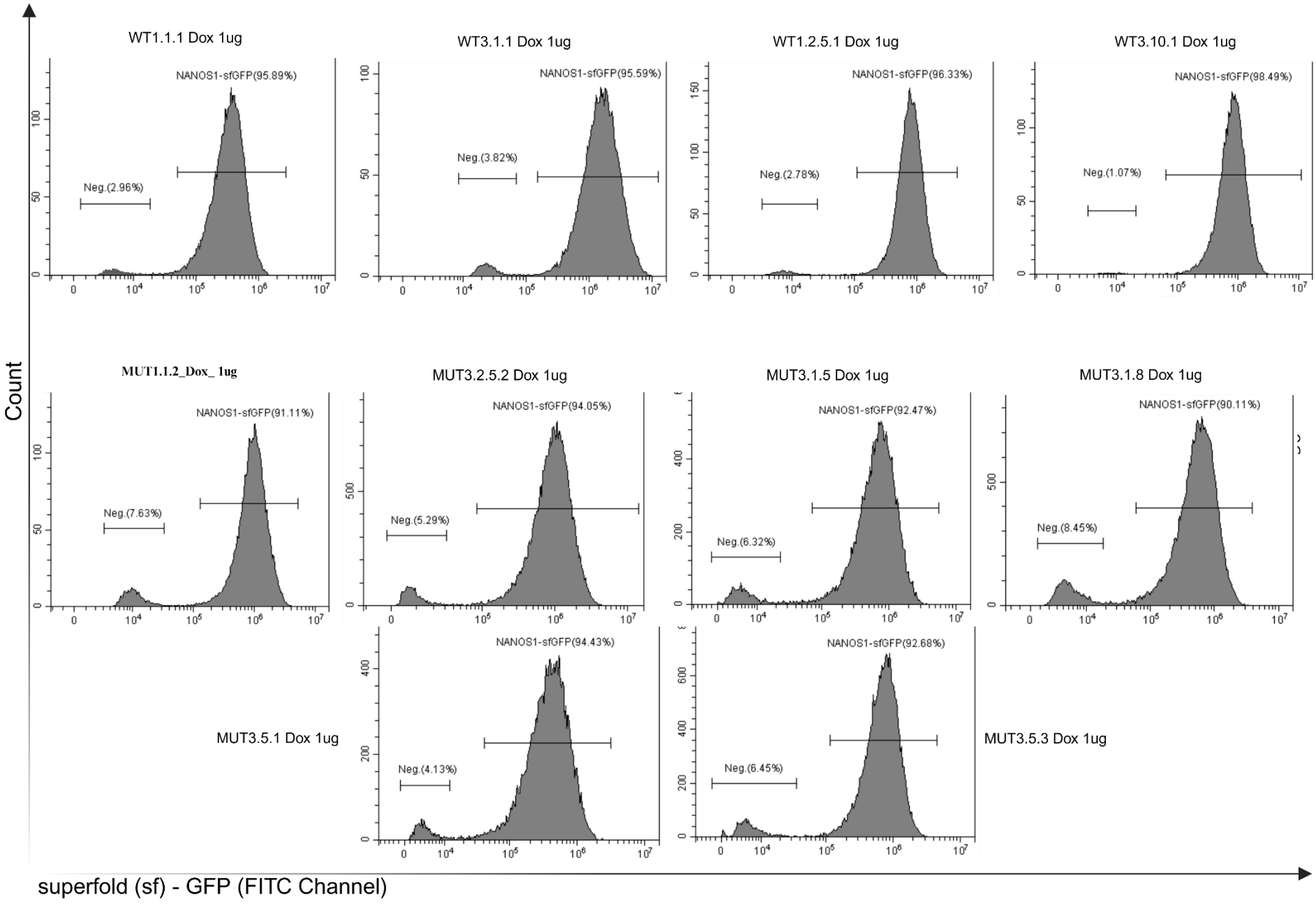
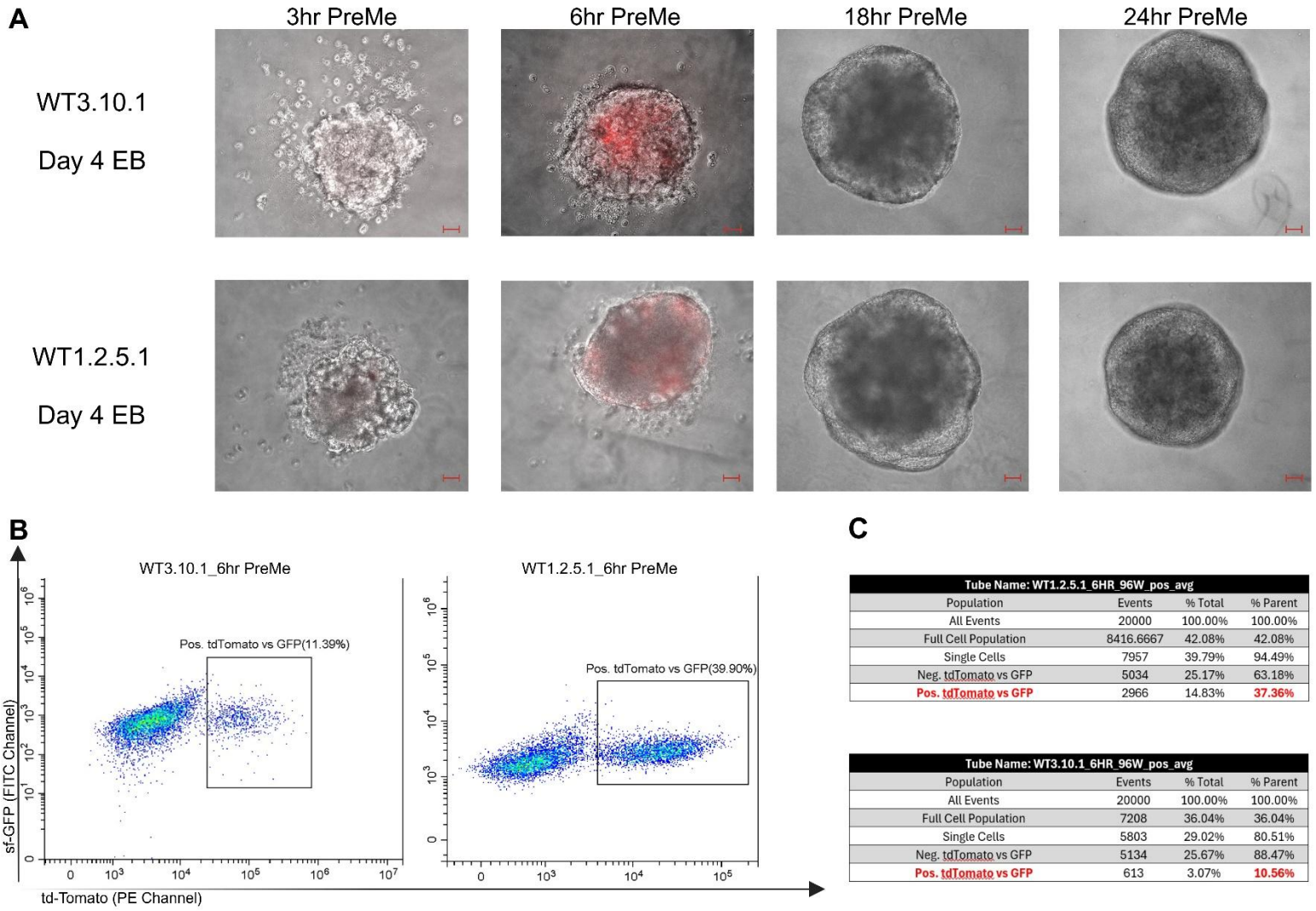

WT1.1.1 Dox 1ug
NANOS1-sfGFP(95.89%)
Neg.(2.96%)
WT3.1.1 Dox 1ug
NANOS1-sfGFP(95.59%)
Neg.(3.82%)
WT1.2.5.1 Dox 1ug
NANOS1-sfGFP(96.33%)
Neg.(2.78%)
WT3.10.1 Dox 1ug
NANOS1-sfGFP(98.49%)
Neg.(1.07%)
MUT1.1.2_Dox_ 1ug
NANOS1-sfGFP(91.11%)
Neg.(7.63%)
MUT3.2.5.2 Dox 1ug
NANOS1-sfGFP(94.05%)
Neg.(5.29%)
MUT3.1.5 Dox 1ug
NANOS1-sfGFP(92.47%)
Neg.(6.32%)
MUT3.1.8 Dox 1ug
NANOS1-sfGFP(90.11%)
Neg.(8.45%)
MUT3.5.1 Dox 1ug
NANOS1-sfGFP(94.43%)
Neg.(4.13%)
MUT3.5.3 Dox 1ug
NANOS1-sfGFP(92.68%)
Neg.(6.45%)
Count
superfold (sf) - GFP (FITC Channel)

Figure 11. Histograms illustrate sfGFP-NANOS1 overexpression (right peak) in the generated PiggyBac cell lines. Following induction with dox (1 µg/ml), fluorescence levels of sfGFP-NANOS1 were quantified using the FITC channel (481 nm wavelength). The x-axis represents fluorescence intensity, while the y-axis denotes the frequency of occurrence (i.e. the number of cells that exhibit fluorescence intensities within each defined range (bin) along the x-axis). A total of 10,000 human embryonic stem cells per sample were analyzed to assess fluorescence intensity and quantify sfGFP-NANOS1 expression.

Table 2. WT- and MUT- NANOS1 PiggyBac cell lines generated through FACS and validated via flow cytometry. *The percentage of sfGFP-positive cells was determined after a 24 h incubation with 1 µg/ml dox. Stem cell morphology prior to induction was also assessed, and all cells were collected at the stem cell stage.

FACS Seeded Clones	Wild-type vs Double Variant	Single cell vs 3-cell Seeding	Transposase: Transposon Ratio	% sfGFP-emitting cells*
WT 1.1.1	Wild-Type	Single Cell	1:1	95.89%
WT 3.1.1	Wild-Type	3-Cell	1:1	95.90%
WT 1.2.5.1	Wild-Type	Single Cell	1:2.5	96.45%
WT 3.10.1	Wild-Type	3-Cell	1:10	98.35%
MUT 1.1.2	Variant	Single Cell	1:1	91.82%
MUT 3.1.5	Variant	3-Cell	1:1	91.25%
MUT 3.1.8	Variant	3-Cell	1:1	90.00%
MUT 3.2.5.2	Variant	3-Cell	1:2.5	94.05%
MUT 3.5.1	Variant	3-Cell	1:5	93.50%
MUT 3.5.3	Variant	3-Cell	1:5	92.68%

Two WT-NANOS1 cell lines with the highest percentage of sfGFP-expressing cells, (1.2.5.1, 96.45%) and (3.10.1, 98.35%) (**Tab. 2**), were chosen for further differentiation experiments aimed at generating PGCs. At the 12 h pre-me time point established for the unedited cell line, neither of the edited cell lines underwent differentiation (data not shown). Consequently, earlier (3 and 6 h) and later (18 and 24 h) pre-me time points were tested, as outlined in **Fig. 12A**. Initial microscopy analysis on Day 4 EBs indicated that for both WT-NANOS1 1.2.5.1 and 3.10.1, the most favorable pre-me window was at 6 h, as the strong tdTomato signal was observed (**Fig. 12A**). The EBs from Day 4 of differentiation were subsequently digested and analyzed *via* flow cytometry, which detected a population of tdTomato-positive PGCs in both cell lines (**Fig. 12B**). On average, WT-NANOS1 cell line 1.2.5.1 exhibited a higher differentiation efficiency at 37.36%, compared to 10.56% for 3.10.1 cell line (**Fig. 12C**). This was further supported by the number of cells gated for the positive PGCLC population, with WT-NANOS1 1.2.5.1 showing 2 966 cells versus 613 cells for 3.10.1 cell line (**Fig. 12C**).



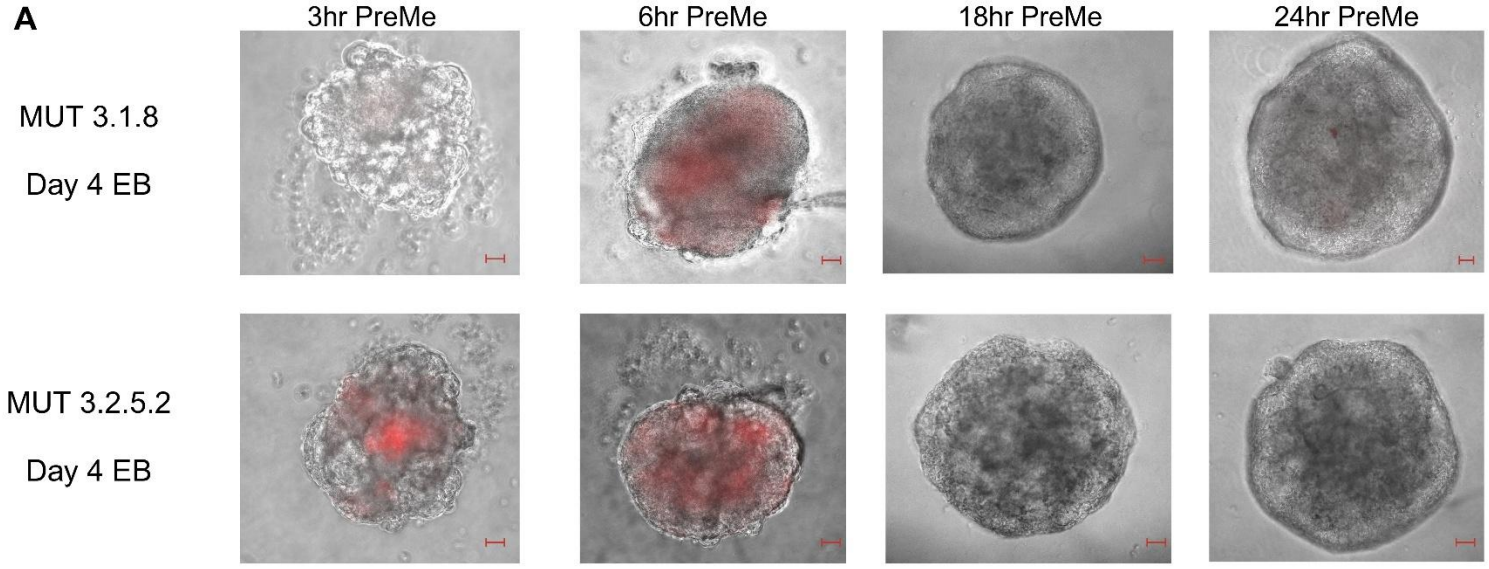
Tube Name: WT1.2.5.1_6HR_96W_pos_avg			
Population	Events	% Total	% Parent
All Events	20000	100.00%	100.00%
Full Cell Population	8416.6667	42.08%	42.08%
Single Cells	7957	39.79%	94.49%
Neg. tdTomato vs GFP	5034	25.17%	63.18%
Pos. tdTomato vs GFP	2966	14.83%	**37.36%**

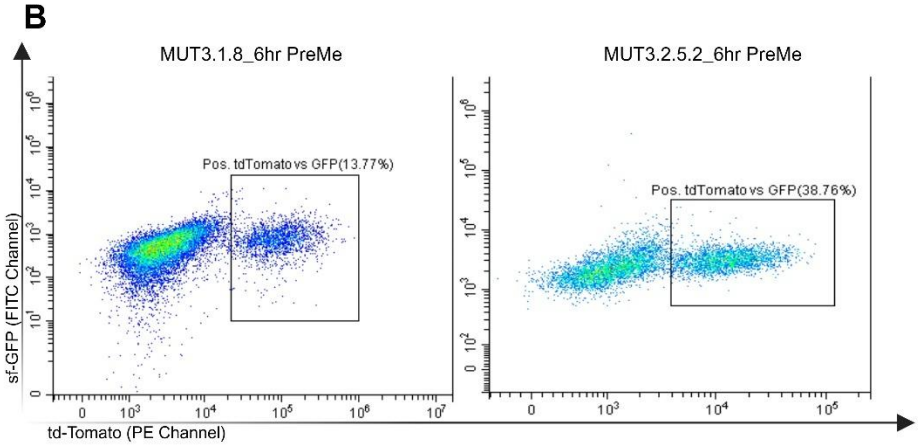
Tube Name: WT3.10.1_6HR_96W_pos_avg			
Population	Events	% Total	% Parent
All Events	20000	100.00%	100.00%
Full Cell Population	7208	36.04%	36.04%
Single Cells	5803	29.02%	80.51%
Neg. tdTomato vs GFP	5134	25.67%	88.47%
Pos. tdTomato vs GFP	613	3.07%	**10.56%**

Figure 12. Differentiation of two WT-NANOS1 1.2.5.1 and 3.10.1 cell lines into PGCs. A- Microscopy images of Day 4 EBs formed at different pre-me time points: 3, 6, 18 and 24 h. tdTomato fluorescence was captured at 561 nm. Scale bar: 20 µm. **B**- Flow cytometry analysis of EBs from two WT-NANOS1 edited cell lines 1.2.5.1 and 3.10.1 at 6 h pre-me, with PGC populations gated using the tdTomato (PE channel) and GFP (FITC channel) signals. **C**- Compiled flow cytometry data from three independent differentiation experiments at 6 h pre-me for the two edited cell lines, 1.2.5.1 (top) and 3.10.1 (bottom). Highlighted in bold red indicates gating of tdTomato-positive (PE channel) versus GFP-positive (FITC channel) cells, representing the average percentage of the PGC population. For both 1.2.5.1 and 3.10.1 cell lines 20,000 digested EB cells were analyzed per differentiation. Gating was applied to identify the target population, single-cell events, and tdTomato-negative versus tdTomato-positive populations prior to dox treatment.

Among the MUT-NANOS1 cell lines, those with the highest levels of sfGFP-NANOS1 overexpression were initially selected for differentiation trials, including 3.2.5.2 and 3.5.1, which showed transgene expression rates of 94.05% and 93.5%, respectively (**Tab. 2**). Similar to the WT-NANOS1 cell lines, the MUT-NANOS1 cell lines did not undergo differentiation at the 12 h pre-me time point (data not shown). However, cell line 3.2.5.2 responded to cytokine treatment after differentiation toward EBs at the 6 h pre-me mark (**Fig. 13A**), whereas cell line 3.5.1 failed to respond at any tested time points (data not shown). Further testing of

the remaining MUT-NANOS1 edited cell lines (**Tab. 2**) revealed that only 3.1.8 exhibited a similar pattern to cell line 3.2.5.2, with its Day 4 EBs expressing the tdTomato signal at satisfactory levels after the 6 h pre-me treatment (**Fig. 13A**). Notably, MUT-NANOS1 cell line 3.1.8 displayed a more complete EB morphology at 6 h compared to the less concluded morphology observed at the 3 h pre-me time point despite the tdTomato signal being apparent also in this timeframe (**Fig. 13A**). Flow cytometry analysis of both successfully differentiated MUT-NANOS1 cell lines indicated that 3.2.5.2 had a higher differentiation efficiency of 38.76%, compared to 13.77% for 3.1.8 (**Fig. 13B**). These findings were consistent across three independent differentiation experiments (**Fig. 13C**). Overall, the average population of PGCs was higher in 3.2.5.2 at 36.26%, compared to 13.13% for 3.1.8, although direct comparison of derived cell numbers was inconclusive due to differences in starting populations and gating strategies (**Fig. 13C**).





C

MUT_3.1.8_6HR_96W_pos_avg			
Population	Events	% Total	% Parent
All Events	13757.5	100.00%	100.00%
Full Cell Population	9689	69.51%	69.51%
Single Cells	9598	68.89%	99.11%
Neg. tdTomato vs GFP	7844	56.33%	81.77%
Pos. tdTomato vs GFP	1289.5	9.06%	**13.13%**

MUT_3.2.5.2_6HR_96W_pos_avg			
Population	Events	% Total	% Parent
All Events	20000	100.00%	100.00%
Full Cell Population	8686	43.43%	43.43%
Single Cells	7551	37.76%	88.47%
Neg. tdTomato vs GFP	4736	23.68%	62.94%
Pos. tdTomato vs GFP	2718	13.59%	**36.26%**

Figure 13. Differentiation of two MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 into PGCs. **A**- Microscopy images of Day 4 EBs formed at different pre-me time points: 3, 6, 18 and 24 h. tdTomato fluorescence was captured at 561 nm. Scale bar: 20 µm, except for cell line 3.1.8 at 24 h pre-me, where the scale bar is 10 µm. **B**- Flow cytometry analysis of digested EBs from two edited cell lines 3.1.8 and 3.2.5.2 at 6 h pre-me, with PGC populations gated using the tdTomato (PE channel) and GFP (FITC channel) signals. **C**- Compiled flow cytometry data from three independent differentiation experiments at 6 h pre-me for the two MUT-NANOS1 cell lines. Highlighted in bold red indicates gating of tdTomato-

positive (PE channel) versus GFP-positive (FITC channel) cells, representing the average percentage of the PGC population. For 3.2.5.2, 20 000 digested EB cells were analyzed per differentiation, while 13 757 cells were analyzed for cell line 3.1.8. Gating was applied to identify the target population, single-cell events, and tdTomato-negative versus tdTomato-positive populations prior to doxycycline treatment.

3.2.1 Dox addition trial in WT- and MUT- NANOS1 cell lines in obtaining PGCs

After successfully establishing differentiation protocols for selected WT- and MUT-NANOS1 cell lines prior to dox addition, trials were conducted to determine the optimal timing within the 4-day differentiation protocol and the appropriate dox concentration, with the goal of inducing overexpression of the NANOS1 cassette without hindering differentiation or sample collection of the PGCs (Day 4). To achieve this, dox addition experiments (**Fig. 14**) were carried out over a 4-day differentiation protocol. Initially, two dox concentrations were tested on the unedited control cell line W15, where the most significant increase in PGC numbers was observed when dox was added on Day 3. In contrast, lower differentiation efficiency was seen when dox was added on Day 0 or Day 2, although the differences on efficiency percentages between concentrations were less pronounced (**Fig. 14**).

In the cell lines WT-NANOS1 1.2.5.1 and MUT-NANOS1 3.2.5.2 selected for further functional studies, the addition of dox on Day 0 resulted in a pronounced inhibitory effect on differentiation at both concentrations tested (200 ng/ml and 500 ng/ml) (**Fig. 14**). Specifically, in the WT-NANOS1 cell line, differentiation was almost completely suppressed. When dox was added on Day 2, the results showed slight improvement but remained inconsistent across both WT- and MUT-NANOS1 cell lines (**Fig. 14**). For instance, in WT-NANOS1 1.2.5.1, a high degree of variability was observed among replicates at the 200 ng/ml concentration, suggesting incomplete dox penetrance. At 500 ng/ml, differentiation efficiency was markedly reduced, with an average of only 2% tdTomato-positive cells. Similarly, MUT-NANOS1 3.2.5.2, the 200 ng/ml concentration also resulted in high variability among replicates, further indicating uneven dox penetrance. However, at 500 ng/ml, differentiation efficiency improved, and variability among replicates was reduced (**Fig. 14**). The most reliable outcomes were achieved when dox was added on Day 3 of differentiation, one day before EB digestion and analysis. At this time point, 500 ng/ml of dox significantly enhanced differentiation efficiency in both cell lines, compared to 200 ng/ml, in which penetrance issues persisted due to the more compact structure of the EBs at this advanced differentiation stage (**Fig. 14**). For the double-positive population representing NANOS1-overexpressing PGCLCs, 500 ng/ml consistently yielded higher cell numbers and lower variability among replicates. Given the overall trend of improved differentiation efficiency with Day 3 dox addition at the higher concentration across representative WT- and MUT- NANOS1 cell lines tested, this protocol was selected for subsequent analysis and sample collection.

As anticipated, early dox induced NANOS1 overexpression or simply dox presence in the culture media disrupted the initial stages of differentiation when cells were still adapting to the cytokine-rich differentiation media. By Day 3, however, the EBs had sufficiently incubated to establish proper differentiation pathways, enabling the generation of the desired PGC population without interference. Thus,

adding dox at this stage allows for a focused assessment of NANOS1 overexpression effects, without confounding disruptions caused by dox or simply protein overexpression during earlier differentiation stages.

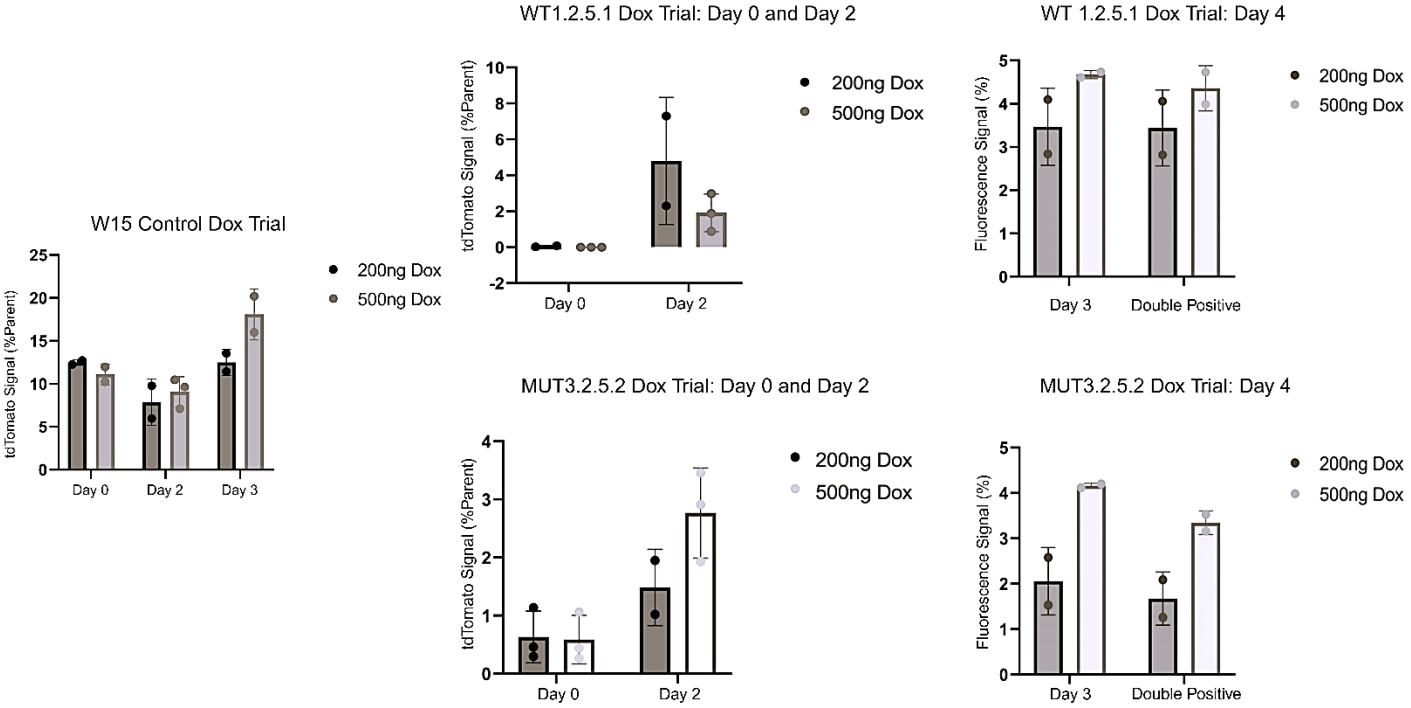


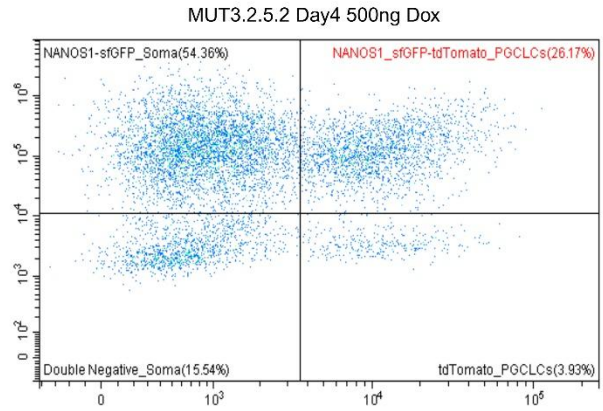
Figure 14. Optimization trials to determine the optimal concentration and timing of dox addition during PGC differentiation. The cell lines used in the dox trials include WT-NANOS1 1.2.5.1 and MUT-NANOS1 3.2.5.2 edited cell lines, as well as the W15 hESC control cell line. Dox was added at three distinct time points: Day 0 (immediately following pre-me incubation), Day 2 (two days following PGC cytokine cocktail incubation), and Day 3 (one day prior to EB digestion and analysis) in the context of a 4-day differentiation protocol. Dox concentrations ranged from 200 ng/ml to 500 ng/ml for both the unedited W15 control and NANOS1 cell lines (left panel for the unedited W15 control; top panel for WT-NANOS1 cell line and bottom panel for MUT-NANOS1 cell line). The experiments were conducted across three independent differentiation trials, with outlier replicates excluded to reduce deviation. The y-axis indicates the percentage of tdTomato-expressing cells gated from the single-cell parent population. In the Day 3 graphs (top and bottom), the double-positive population represents sfGFP-NANOS1 and tdTomato-expressing PGCs, with fluorescence signals corrected *via* flow cytometry compensation.

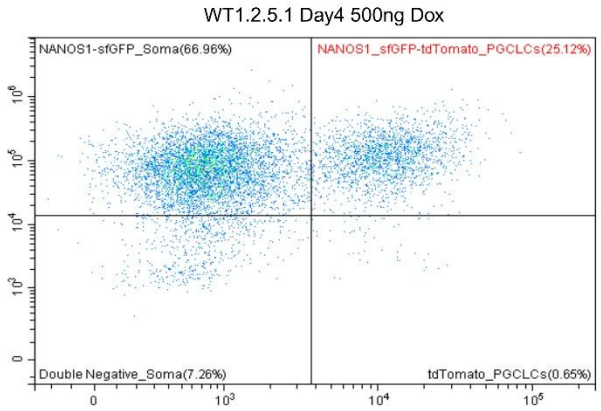
3.2.2 Dox-induced WT-NANOS1 (1.2.5.1) and MUT-NANOS1 (3.2.5.2) cell lines successfully generate sufficient PGCs

The two highest-performing NANOS1 cell lines, WT 1.2.5.1 (**Fig. 12A-C**) and MUT 3.2.5.2 (**Fig. 13A-C**), were successfully differentiated under NANOS1 overexpression conditions, following the addition of 500 ng/ml dox on Day 3 of the 4-day differentiation protocol directed toward PGCs (**Fig. 15A-C**). This strategy was chosen based on previous dox addition trials (**Fig. 14**), which demonstrated that Day 3 was the optimal time point for dox induction, allowing for efficient NANOS1 overexpression without negatively affecting PGC differentiation process. Flow cytometry analysis revealed that the proportion of double-positive (sfGFP-tdTomato) PGCLCs overexpressing NANOS1 was 26.17% in MUT 3.2.5.2 and 25.12% in the counterpart WT 1.2.5.1 PGLCs (**Fig. 15A**). Across three independent differentiation

experiments, the average population of germ cells overexpressing NANOS1 was 31.88% for MUT 3.2.5.2 and 27.15% for WT 1.2.5.1, therefore relatively comparable and stable (**Fig. 15B**). Moreover, microscopy analysis demonstrated a well-defined EB structure in both cell lines, with clear co-expression of sfGFP from the NANOS1 cassette and tdTomato fluorescence highlighting the emerging PGCLC population driven by NANOS3 expression (**Fig. 15C**).

A



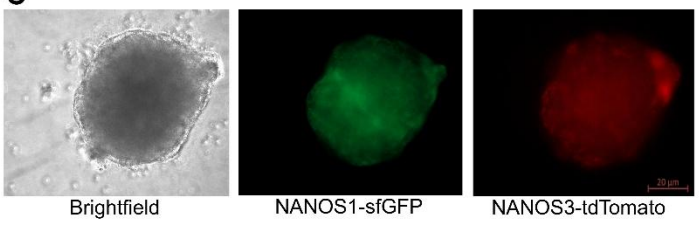


B

MUT3.2.5.2_NANOS1 PGCLCs Day 4_500ng Doxycycline			
Population	Events	% Total	% Parent
All Events	20000	100.00%	100.00%
Full Population	7383	36.92%	36.92%
Single Cells	6982	34.91%	94.74%
NANOS1_sfGFP-tdTomato_PGCLCs	2206	11.03%	31.88%
NANOS1-sfGFP_Soma	3593	17.97%	51.31%
Double Negative_Soma	937	4.69%	13.32%
tdTomato_PGCLCs	246	1.23%	3.50%

WT 1.2.5.1_NANOS1 PGCLCs Day 4_500ng Doxycycline			
Population	Events	% Total	% Parent
All Events	20000	100.00%	100.00%
Full Population	7576	37.88%	37.88%
Single Cells	6914	34.57%	91.56%
NANOS1_sfGFP-tdTomato_PGCLCs	1899	9.50%	27.15%
NANOS1-sfGFP_Soma	4557	22.79%	66.13%
Double Negative_Soma	404	2.02%	6.01%
tdTomato_PGCLCs	53	0.27%	0.72%

C



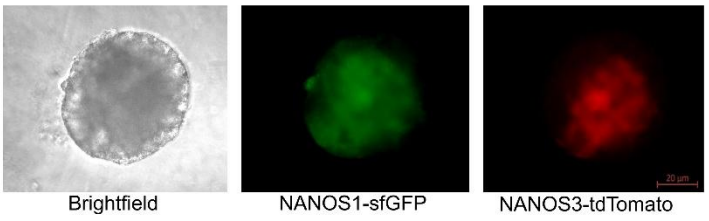


Figure 15. Analysis of WT- and MUT-NANOS1 cell lines following dox induction, assessing differentiation efficiency toward PGCLCs characterized by respective NANOS1 overexpression. A- Flow cytometry dot plots for MUT (left) and WT (right), showing cell populations derived from Day 4 EBs after the addition of 500 ng/ml dox on Day 3 of differentiation. The double-positive population for NANOS1 (sfGFP) and PGCLCs (tdTomato) is highlighted in red, with fluorescence detected in the FITC (y-axis) and PE (x-axis) channels. **B**- Compiled flow cytometry data from three independent differentiation experiments for MUT (left) and WT (right). Gating was applied to 20,000 digested EB cells, distinguishing four key populations: NANOS1 (sfGFP) and PGCLCs (tdTomato) double-positive cells (highlighted in red), NANOS1 (sfGFP)-expressing somatic cells, double-negative somatic cells, and tdTomato-only PGCLCs. **C**- Microscopy images of Day 4 EBs for MUT (left) and WT (right), captured in brightfield, with fluorescence from NANOS1-sfGFP (510 nm emission) and PGCLCs expressing NANOS3-tdTomato (581 nm emission). Scale bar: 20 µm.

3.2.3 Validated NANOS1 protein overexpression in WT-NANOS1 and MUT-NANOS1 cell lines following dox induction

Furthermore, validation of NANOS1 cassette protein expression using an antibody targeting the 3x-FLAG tag conjugated to NANOS1 (**Fig. 16B-C**) demonstrated precise regulation by the Tet-ON system, with no detectable leakage of cassette expression in the absence of dox induction. The specificity of the antibody for the introduced cassette was confirmed, as no signal was observed in the unedited W15 cell line (W15 Ctrl, **Fig. 16B-C**). Protein loading appeared consistent across the samples, as indicated by housekeeping (vinculin) protein levels, except for the MUT-NANOS1 3.1.8 cell line, where slightly elevated starting protein levels were added for both its before and after dox treated samples relative to the other cell lines (**Fig. 16A**). In terms of NANOS1-3xFLAG-P2A protein band intensity, the highest levels were observed in WT 1.2.5.1 following dox induction, followed by MUT 3.1.8 (**Fig. 16B-C**). The elevated band intensity in MUT 3.1.8 is likely attributable to the increased amount of initial protein loaded compared to the other samples (**Fig. 16A**). Therefore, the NANOS1 protein validation assay *via* western blot served as further support for the selection of cell lines WT 1.2.5.1 and MUT 3.2.5.2 as the most suitable for additional high-throughput analysis, due to their comparable and reliable levels of NANOS1 protein overexpression (**Fig. 16B-C**).

These results provide robust evidence for the ability of the generated cell lines to successfully differentiate before and after NANOS1 overexpression, yielding sufficient cell numbers for subsequent characterization and analysis. Moreover, the stringent control of NANOS1 expression by the Tet-ON system prior to dox induction confirms that the dox-negative replicates are appropriate for further validation and analysis. The two best-performing cell lines, MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1, were identified as optimal candidates for high-throughput analysis, while the remaining top-performing clones, MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1, were designated as suitable candidates for validation assays to mitigate potential cell line variation when assessing targets of interest.

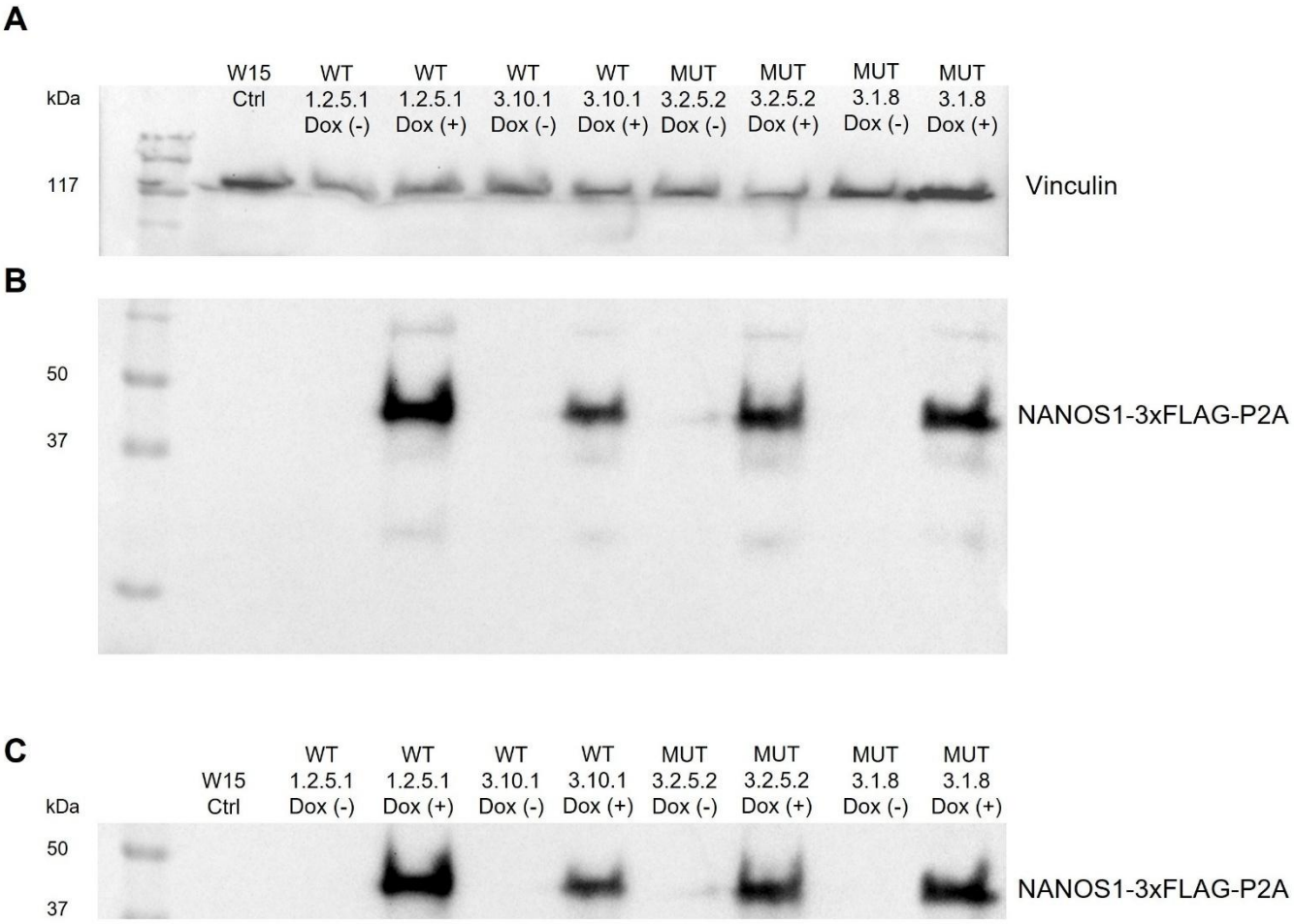


Figure 16. Western blot analysis validating the overexpression of NANOS1-3xFLAG in WT-NANOS1 (1.2.5.1 and 3.10.1) and MUT-NANOS1 (3.2.5.2 and 3.1.8) cell lines. The cells were collected at the pre-me stage of differentiation before and after dox addition. **A**- Reference blot for the housekeeping protein vinculin (117 kDa; 1:10,000 dilution) with a 3'' exposure, shown for the control cell line (W15 Ctrl) and remaining cell lines, both pre- and post-dox treatment. **B**- Full blot using an anti-FLAG antibody (1:1,500 dilution) with a 3'' exposure to detect the NANOS1-3xFLAG-P2A fusion protein. **C**- Magnified region of the blot between 37 kDa and 50 kDa, focusing on the detection of the NANOS1-3xFLAG-P2A protein using the anti-FLAG antibody.

3.2.4 All selected cell lines (2 WT- and 2 MUT-NANOS1) exhibit normal 46,XY karyotypes

The NANOS1 cell lines, alongside the human embryonic stem cell (hESC) control line W15 (**Fig. 16A-C),** underwent comprehensive validation via karyotyping and immunofluorescence assays to assess stage-specific markers. As anticipated, the karyotype of the control line W15 was determined to be 46, XY, with confirmation obtained through independent analysis of 10 metaphase spreads. These analyses revealed no chromosomal rearrangements attributable to either cell passaging or *in vitro* handling of the stem cells (**Fig. 17A**). Additionally, the karyotypes of the cell lines, including WT 1.2.5.1 and WT 3.10.1 (**Fig. 17D-E**) as well as MUT 3.2.5.1 and MUT 3.1.8 (**Fig. 17B-C**), were verified to be consistent with W15, displaying a normal 46, XY karyotype. Crucially, none of the cell lines, including W15, exhibited any chromosomal aberrations following PiggyBac transposon-mediated editing, further supporting their genomic stability and confirming their appropriateness for subsequent experimental analyses.

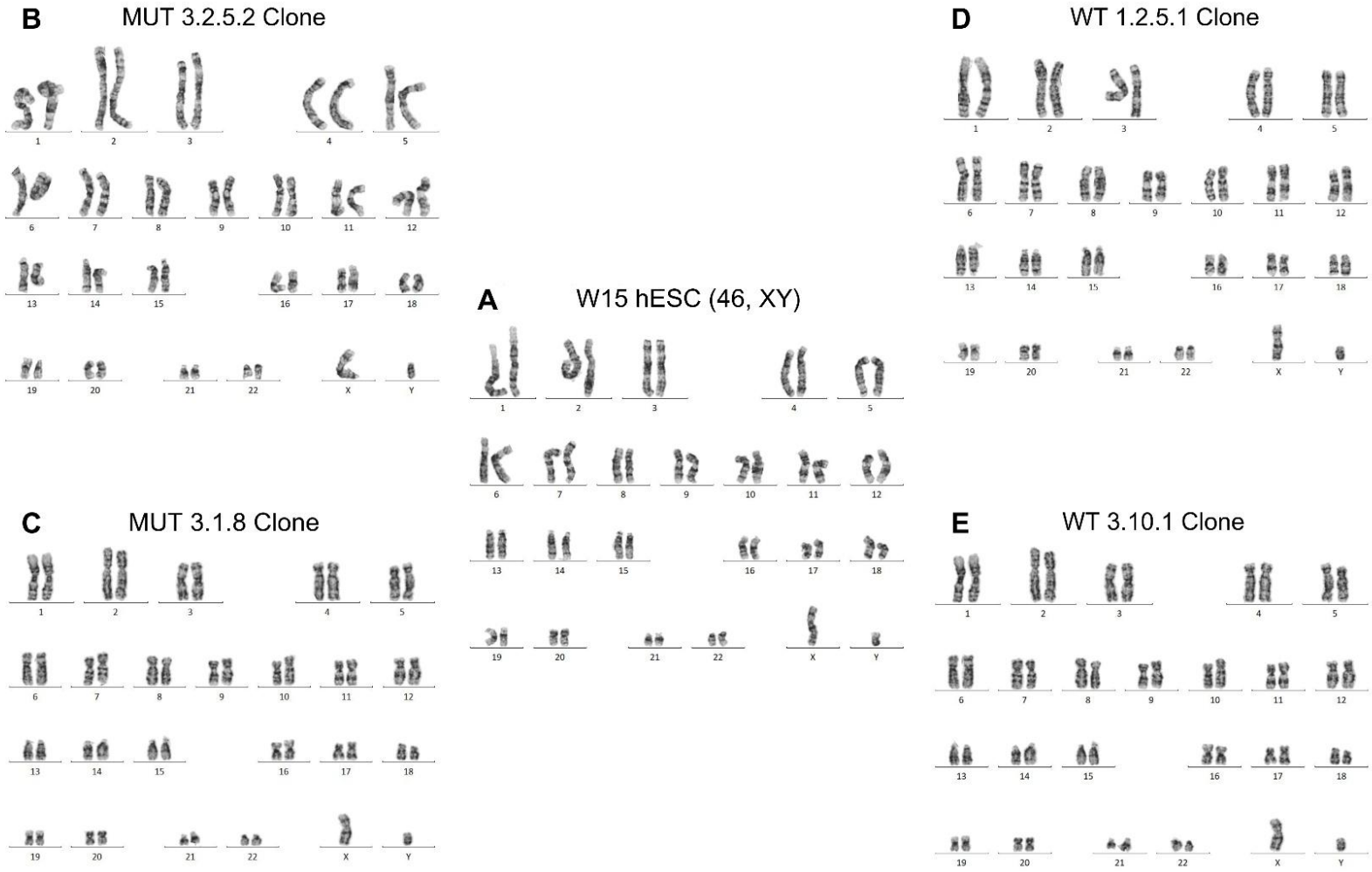

B
MUT 3.2.5.2 Clone
A
W15 hESC (46, XY)
D
WT 1.2.5.1 Clone
C
MUT 3.1.8 Clone
E
WT 3.10.1 Clone

Figure 17. Karyotype analysis of 10 metaphases for selected cell lines and the hESC control cell line W15. **A**- Karyotype of the unedited hESC line W15, showing 22 autosomal chromosomes and XY sex chromosomes, confirming a normal karyotype of 46, XY. **B-C** Karyotypes of MUT NANOS1 3.2.5.2 and 3.1.8 cell lines, both displaying a normal karyotype of 46, XY. **D-E** karyotypes of WT NANOS1 1.2.5.1 and 3.10.1 cell lines, both showing a normal karyotype of 46, XY.

3.2.5 All selected cell lines (2 WT- and 2 MUT-NANOS1) express appropriate nuclear and cytoplasmic stem cell markers

The initial assessment of the cell lines was conducted at the stem cell stage, where the expression of key pluripotency markers, including NANOG, OCT4, and SOX2, was evaluated. These markers are critical indicators of the cells' ability to retain their pluripotent state, which is essential for stem cells to differentiate into multiple cell types. As illustrated in **Fig. 18**, all the tested cell lines demonstrated expression of these pluripotency factors, albeit at varying levels, which supports the conclusion that the generated cell lines maintain characteristics of stemness and pluripotency similar to the control line W15. The differential expression observed among the cell lines provides further insight into their relative pluripotent potential.

Specifically, the MUT-NANOS1 cell lines 3.2.5.2 and 3.1.8 exhibited higher expression levels of NANOG, OCT4, and SOX2 compared to their WT-NANOS1 counterparts 1.2.5.1 and 3.10.1. This elevated expression in the MUT-NANOS1 cell lines suggests an enhanced maintenance of pluripotency, indicating that these cell lines may have a greater potential for differentiation into PGC-like cells *in vitro*. The results were obtained without the addition of dox, ensuring that the cell lines were in an uninduced state, which provides a clearer comparison of their intrinsic pluripotent capacities. The qualitative differences in pluripotency between the cell lines were most notable in the WT-NANOS1 cell lines with 1.2.5.1 exhibiting the lowest expression of the pluripotency markers, indicating a reduced capacity to retain its undifferentiated state. Instead, the WT 3.10.1 cell line showed a moderate level of pluripotency, falling between WT-NANOS1 1.2.5.1 and the MUT-NANOS1 cell lines. In contrast, the MUT cell lines, 3.2.5.2 and 3.1.8, demonstrated the highest levels of pluripotency marker expression, suggesting that they are more robust in maintaining an undifferentiated, pluripotent state. These findings suggest that the MUT-NANOS1 cell lines are qualitatively more pluripotent than the WT-NANOS1 counterpart, particularly at the uninduced stem cell stage. Moreover, the apparent hierarchy of pluripotency, with WT-NANOS1 1.2.5.1 being the least pluripotent and MUT-NANOS1 3.1.8 being the most pluripotent, highlights the significant impact of genetic editing on the stem cell state. These differential expression patterns among the cell lines provide a valuable foundation for interpreting further results obtained through high-throughput sequencing.

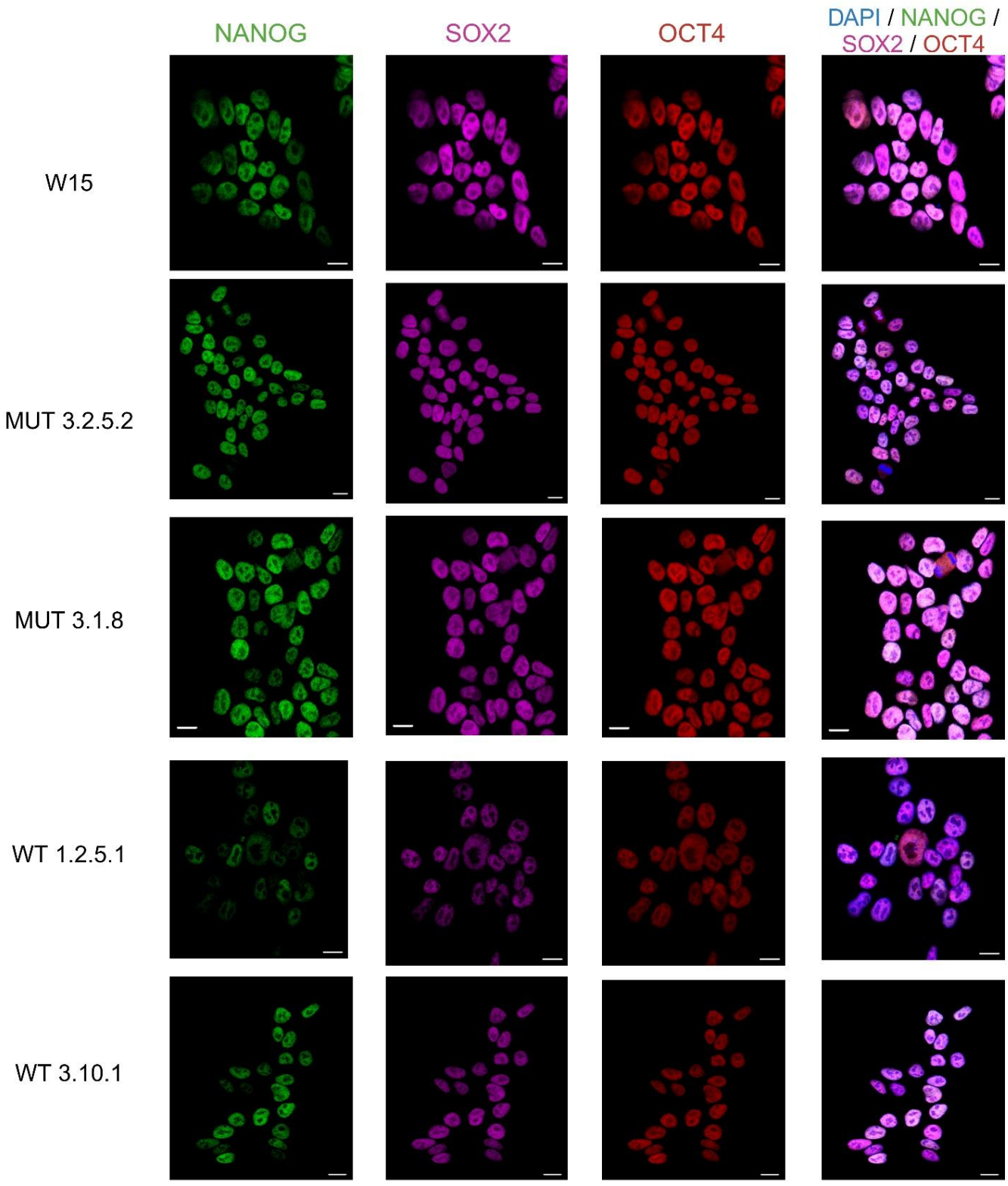


Figure 18. **Immunofluorescence analysis of stem cell markers NANOG, SOX2, and OCT4 in the hESC control line W15 and edited cell lines (monolayer).** MUT-NANOS1 cell lines 3.2.5.2 and 3.1.8, along with WT-NANOS1 cell lines 1.2.5.1 and 3.10.1, were assessed at the stem cell stage while cultured in standard media without dox. Nuclei were stained with DAPI. Scale bar: 50 µm (zoomed in for W15, MUT 3.1.8, and WT 1.2.5.1).

SSEA-4 (Stage-Specific Embryonic Antigen-4) is a key glycosphingolipid marker expressed on the surface of human embryonic stem cells (hESCs) and is widely used to assess the undifferentiated, pluripotent state of stem cells. Its role, while complementary to nuclear pluripotency markers such as NANOG, OCT4, and SOX2, provides a critical indicator of the cell membrane's characteristics, which are often associated with cell surface properties and pluripotency potential. SSEA-4 has been extensively characterized in stem cell research as a reliable marker for identifying hESCs and their ability to remain in a pluripotent, undifferentiated state (Henderson et al., 2002).

In this study, the expression of SSEA-4 was evaluated in all cell lines at the uninduced stage (without dox), adding another dimension to the assessment of their pluripotency. As illustrated in **Fig. 19**, SSEA-4 was detected across all cell lines, mirroring the pattern observed for the nuclear markers NANOG, OCT4, and SOX2 (**Fig. 18**). However, the variability in SSEA-4 expression was less pronounced than that of the nuclear markers. Notably, the WT-NANOS1 1.2.5.1 cell line exhibited slightly reduced SSEA-4 fluorescence intensity (**Fig. 19**), consistent with its lower expression of the nuclear pluripotency markers (**Fig. 18**), which suggests a relatively lower level of stemness for this clone.

In conclusion, while nuclear markers regulate the core transcriptional networks essential for maintaining pluripotency (Takahashi & Yamanaka, 2006), SSEA-4 serves as a complementary cytoplasmic and membrane marker confirming the undifferentiated state. The observation of SSEA-4 expression in all cell lines validates their classification as pluripotent at the stem cell stage and supports their overall stemness. However, the slight variability in SSEA-4 expression suggests differences in glycosphingolipid surface composition or membrane characteristics among cell lines, potentially influencing their ability to remain undifferentiated or to differentiate. Notably, the reduced expression of both nuclear and cytoplasmic markers in WT-NANOS1 1.2.5.1 indicates that not all cell lines start with the same baseline pluripotency prior to NANOS1 overexpression, which may significantly impact their subsequent differentiation potential or responsiveness to external stimuli.

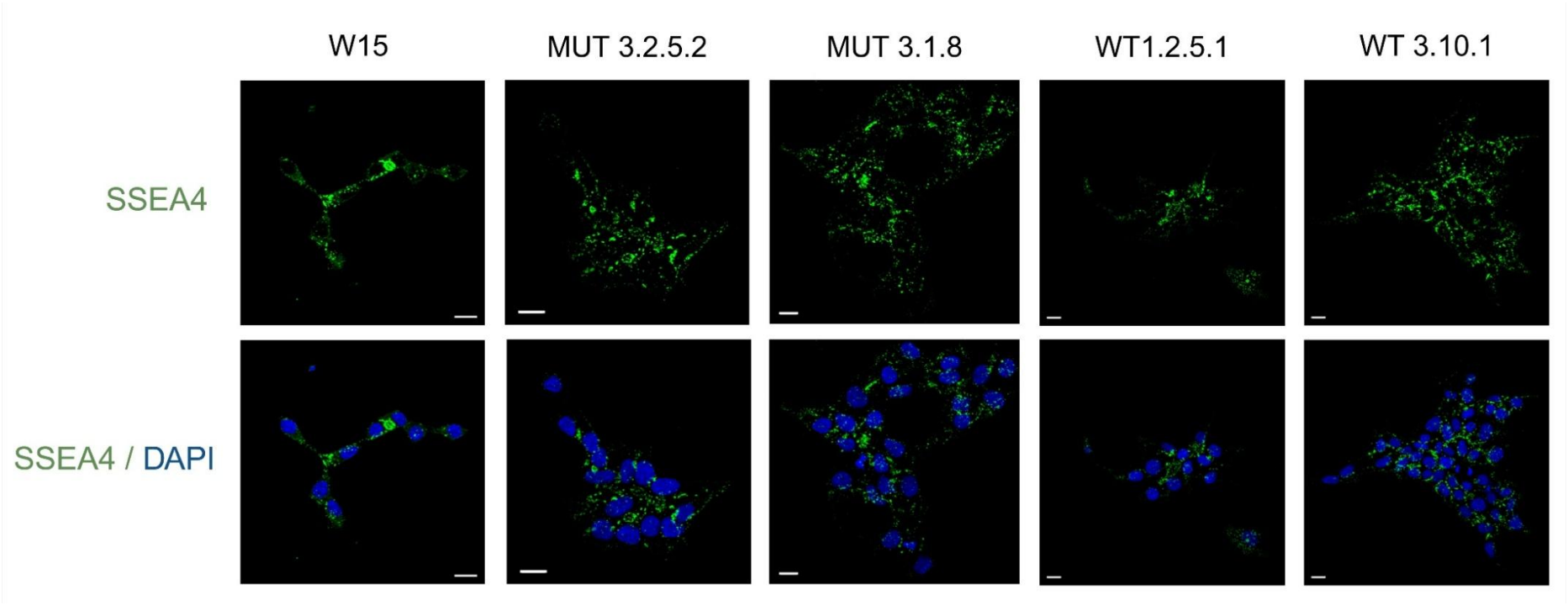

W15
MUT 3.2.5.2
MUT 3.1.8
WT1.2.5.1
WT 3.10.1
SSEA4
SSEA4 / DAPI

Figure 19. Immunofluorescence analysis of the cytoplasmic stem cell marker SSEA-4 in the hESC control line W15 and edited cell lines. MUT-NANOS1 cell lines (3.2.5.2 and 3.1.8) and WT-NANOS1 cell lines (1.2.5.1 and 3.10.1) were assessed at the stem cell stage, cultured as a monolayer under standard media conditions in the absence of dox. Scale bar: 50 μm, with larger magnification provided for W15, MUT-NANOS1 3.2.5.2, and 3.1.8 for enhanced visualization of marker localization and expression levels.

3.2.6 All selected cell lines (2 WT- and 2 MUT-NANOS1) express pre-me markers Brachyury (T) and EOMES

At the pre-me stage, two key markers of PGC competence in hESCs induced *in vitro*, T and EOMES, were closely examined. T is particularly important as it signifies germ cell fate competence thoroughly examined in mice studies. Furthermore, its role in hESCs has also been noted, as expression can be induced by bone morphogenetic proteins (BMPs), making it a critical indicator of PGC potential (Irie et al., 2015). Furthermore, SOX17, a crucial human PGC marker, is expressed in T-positive cells prior to the appearance of PRDM1, marking the initial specification of hPGCLCs (Irie et al., 2015).

In the control cell line W15, T expression was observed both before and after the addition of dox (**Fig. 20B**), suggesting that dox induction does not influence T marker expression in the control hESC line. This stability in T expression indicates that the control cell line retains its competence for PGC fate regardless of dox treatment. In contrast, in the MUT-NANOS1 3.2.5.2 cell line, T-positive cells exhibited a higher level of T fluorescence expression prior to dox addition, but this expression was quantitatively reduced following the overexpression of MUT-NANOS1 after dox induction (**Fig. 20B**). This suggests that overexpression of MUT-NANOS1 may negatively affect the PGC competence marked by the reduced T marker expression in this cell line.

On the other hand, in the WT-NANOS1 1.2.5.1 cell line, the initial expression of T marker was lower compared to the MUT-NANOS1 cell line before dox treatment. Interestingly, the addition of dox and subsequent overexpression of WT-NANOS1 did not result in any noticeable change in T fluorescence intensity, implying that overexpression of WT-NANOS1 has no significant impact on T marker expression in this cell line (**Fig. 20B**). This lack of response may indicate a more stable PGC competence in the WT cell line post-NANOS1 overexpression. Additionally, the expected fluorescence of NANOS1-sfGFP was detected exclusively in the cell lines post-dox addition (**Fig. 20A**), as anticipated. The fluorescence intensity of NANOS1-sfGFP was slightly higher in the WT-NANOS1 1.2.5.1 cell line compared to the MUT-NANOS1 3.2.5.2 cell line, and the protein localization was evident in both the nucleus and cytoplasm of the cell lines (**Fig. 20A**, **Fig. 20C**). This suggests that NANOS1 is effectively overexpressed in both cell lines, with a potentially greater accumulation in the WT-NANOS1 cell line. The differential expression of NANOS1 across generated cell lines and its presence in both nuclear and cytoplasmic compartments may reflect in the post-transcriptional regulation or functional dynamics of NANOS1 in these cells.

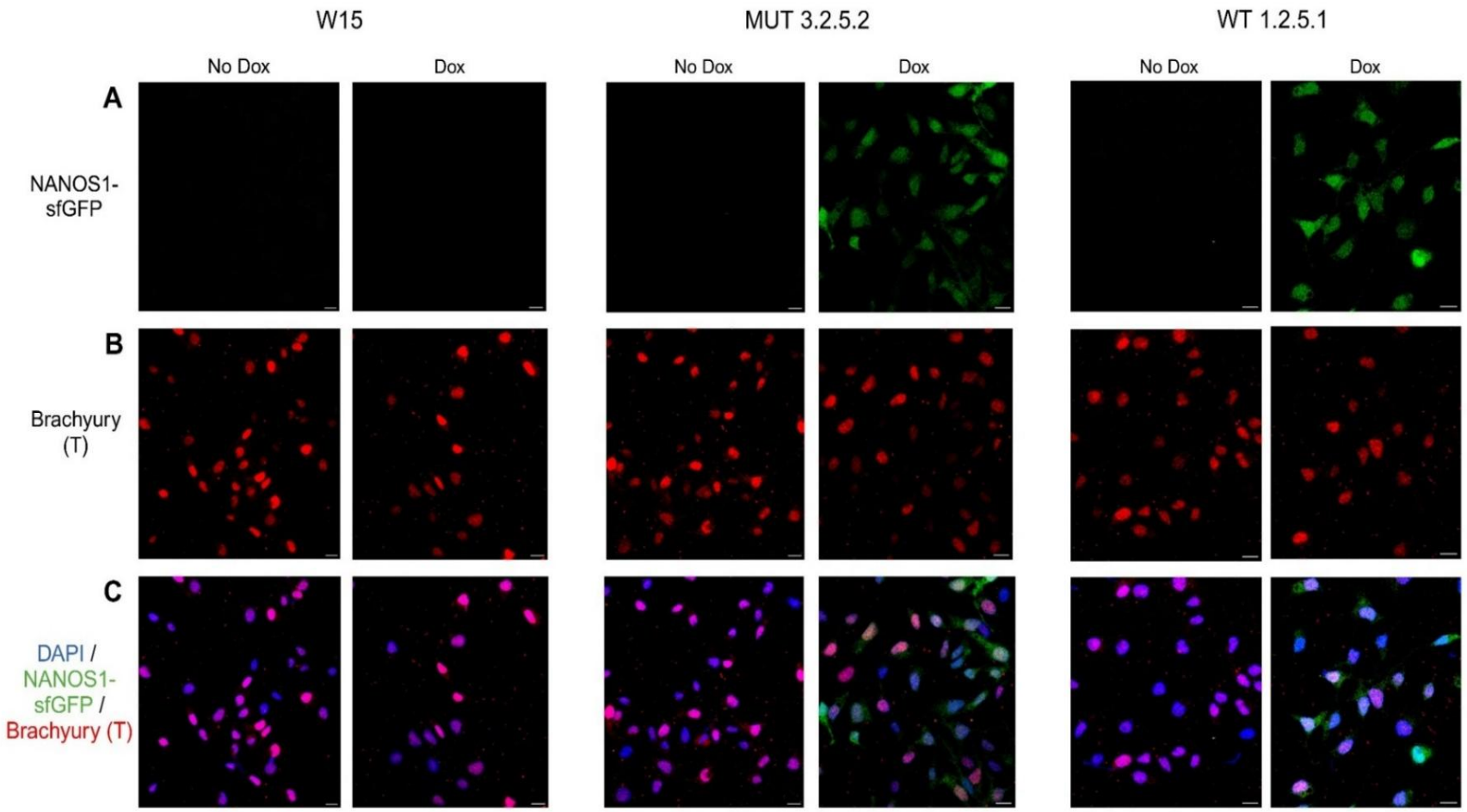

W15
MUT 3.2.5.2
WT 1.2.5.1
No Dox
Dox
No Dox
Dox
No Dox
Dox
A
NANOS1-
sfGFP
B
Brachyury
(T)
C
DAPI /
NANOS1-
sfGFP /
Brachyury (T)

Figure 20 Immunofluorescence analysis of the T marker and NANOS1-sfGFP expression was performed in the hESC control cell line W15 and MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 cell lines. A- NANOS1-sfGFP expression in pre-me cells, fixed 6-hrs post-induction. Dox was added at a concentration of 500 ng/mL at the onset of pre-me induction to initiate NANOS1 overexpression. **B**- Immunostaining for T, a key mesodermal marker indicating the transition toward germ cell competence, was performed on pre-me cells both before and after dox-induced overexpression. **C**- DAPI nuclear staining was used to visualize the nuclei, while the overlap of DAPI, NANOS1-sfGFP, and T fluorescence was assessed across the different cell lines and treatment conditions. Scale bar: 70 μm.

Additionally, the supplementary clones MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 were also examined for the presence of T-positive cells at the pre-me stage. As observed with MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 cell lines, the NANOS1-sfGFP signal was detected only after dox induction, with similar expression levels in both the nucleus and cytoplasm (**Fig. 21A** and **Fig. 21C**). Prior to dox addition, both MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 cell lines exhibited high levels of transcription factor T expression (**Fig. 21B**), a critical marker of pre-me and PGC competence. However, following dox induction, T signal intensity decreased, particularly in MUT-NANOS1 3.1.8, further supporting the notion that suggesting that MUT-NANOS1 overexpression may impair the priming of pre-me cells for human PGC specification by potentially disrupting the pathways necessary to maintain PGC competence.

For the WT-NANOS1 3.10.1 cell line, a similar trend was observed where T marker fluorescence expression levels decreased following the overexpression of WT-NANOS1, unlike in its counterpart WT-NANOS1 1.2.5.1, which did not show a reduction in T signal post-dox induction (**Fig. 20B** and **Fig. 21B**). Before dox treatment, WT-NANOS1 3.10.1 displayed a higher initial T signal, but the subsequent overexpression of WT-NANOS1 led to a significant drop in marker intensity, suggesting that in this supplementary cell line the overexpression of WT-NANOS1 may, under certain conditions, impact the regulation of T expression negatively.

The differences between the WT-NANOS1 cell lines 3.10.1 and 1.2.5.1 and their MUT-NANOS1 counterparts suggest a potential cell line effect. In particular, the overexpression of WT-NANOS1-sfGFP in cell line 3.10.1 was noticeably higher compared to WT-NANOS 1.2.5.1 (**Fig. 21A** vs. **Fig. 20A**). This discrepancy could be linked to the varying ratios of transposon to transposase used during the generation of these cell lines (10:1 for WT-NANOS1 3.10.1 versus 2.5:1 for WT-NANOS1 1.2.5.1; **Tab. 2**), which may have led to a higher level of NANOS1 expression in WT-NANOS1 3.10.1. The increased NANOS1 expression in cell line 3.10.1 may have caused cellular stress or exceeded the optimal levels required for NANOS1 to enhance PGC competence, leading instead to a negative effect on the expression of key markers like transcription factor T.

This suggests that while NANOS1 plays a critical role in PGC specification, there appears to be a threshold for its effective expression. Overexpression beyond this threshold, particularly in WT-NANOS1 3.10.1, could potentially disrupt normal regulatory processes associated with germ cell fate determination. Therefore, the balance of NANOS1 expression is likely crucial in determining the optimal conditions for promoting human PGC specification at the pre-me stage.

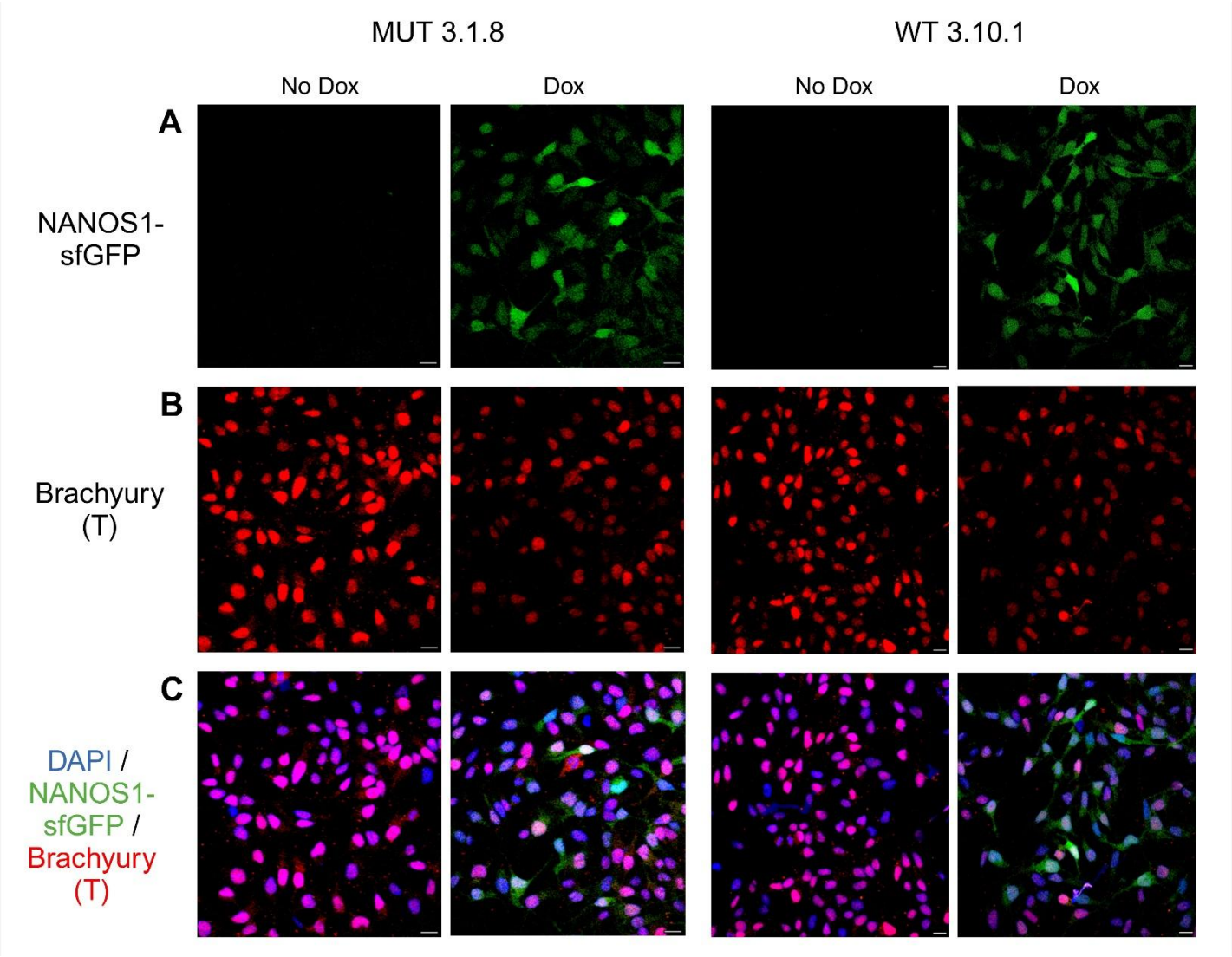


Figure 21. Immunofluorescence analysis of transcription factor T and NANOS1-sfGFP expression in the MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 cell lines to investigate the impact of NANOS1 overexpression on pre-me stage of differentiation. A- Expression of NANOS1-sfGFP in pre-me cells was evaluated 6 h post-induction, with dox added at a concentration of 500 ng/ml at the start of pre-me induction. **B**- Immunostaining for marker T was performed on pre-me cells both before and after dox treatment. The analysis of T expression patterns across both MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 enabled the assessment of how NANOS1 overexpression modulates the cells' commitment to the PGC lineage. **C**- Nuclei were stained with DAPI, and the resulting overlay of DAPI, NANOS1-sfGFP, and T fluorescence provided insights into the spatial localization and co-expression of these markers. Scale bar: 70 µm.

EOMES is another crucial marker for PGC competence during the transition from hESCs to the pre-me stage. During this transition, endogenous fibroblast growth factor (FGF) and WNT signaling pathways activate early mesodermal enhancers and genes, including EOMES, which is essential for hPGC specification (Tang et al., 2022). It has been hypothesized that a temporal increase in EOMES levels plays a pivotal role in defining the acquisition of PGC competence, particularly in the absence of later-activated mesodermal TFs such as Goosecoid homeobox (GSC) and GATA-binding factor 6 (GATA6), which promote somatic differentiation (Tang et al., 2022).

As shown in **Fig. 22B-C**, EOMES was not universally expressed in all detected pre-me cells (nuclei stained with DAPI) in either the control cell line W15 or the cell lines MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1. Similar to TF T, the addition of dox did not significantly affect EOMES expression in the control cell line W15 (**Fig. 22B-C**). However, before dox induction, EOMES levels were lower in MUT-NANOS1 3.1.8 compared to WT-NANOS1 3.10.1, suggesting a cell line effect influencing EOMES levels following genetic editing via PiggyBac (**Fig. 22B-C**).

Following dox induction (**Fig. 22D-F**), NANOS1-sfGFP expression was detected in both cell lines, with WT-NANOS1 3.10.1 showing slightly higher fluorescence intensity (**Fig. 22D**). This elevated expression of NANOS1 in WT-NANOS1 3.10.1 correlated with reduced levels of EOMES, particularly in cells where NANOS1 expression was strongest (**Fig. 22E-F**). This further supports the notion that overexpression of WT-NANOS1 in cell line 3.10.1 may have exceeded the optimal levels, negatively affecting the cells' ability to mediate PGC competence via EOMES and previously TF T.

A similar trend was observed in MUT-NANOS1 3.1.8 cell line after dox induction (**Fig. 22D-F**), where the presence of MUT-NANOS1 also led to a reduction in EOMES expression within the nuclei of affected pre-me cells (**Fig. 22F**). However, the overall EOMES fluorescence intensity in MUT-NANOS1 3.1.8 cell line remained slightly higher than in WT-NANOS1 3.10.1, likely reflecting the cell line differences due to the lower transposon-to-transposase ratio (1:1 versus 10:1; **Tab. 2**) used during the generation of MUT-NANOS1 3.1.8. This difference in transposon integration efficiency may account for the varying impacts on NANOS1 and EOMES expression, highlighting the significance of cell line effects in influencing PGC competence.

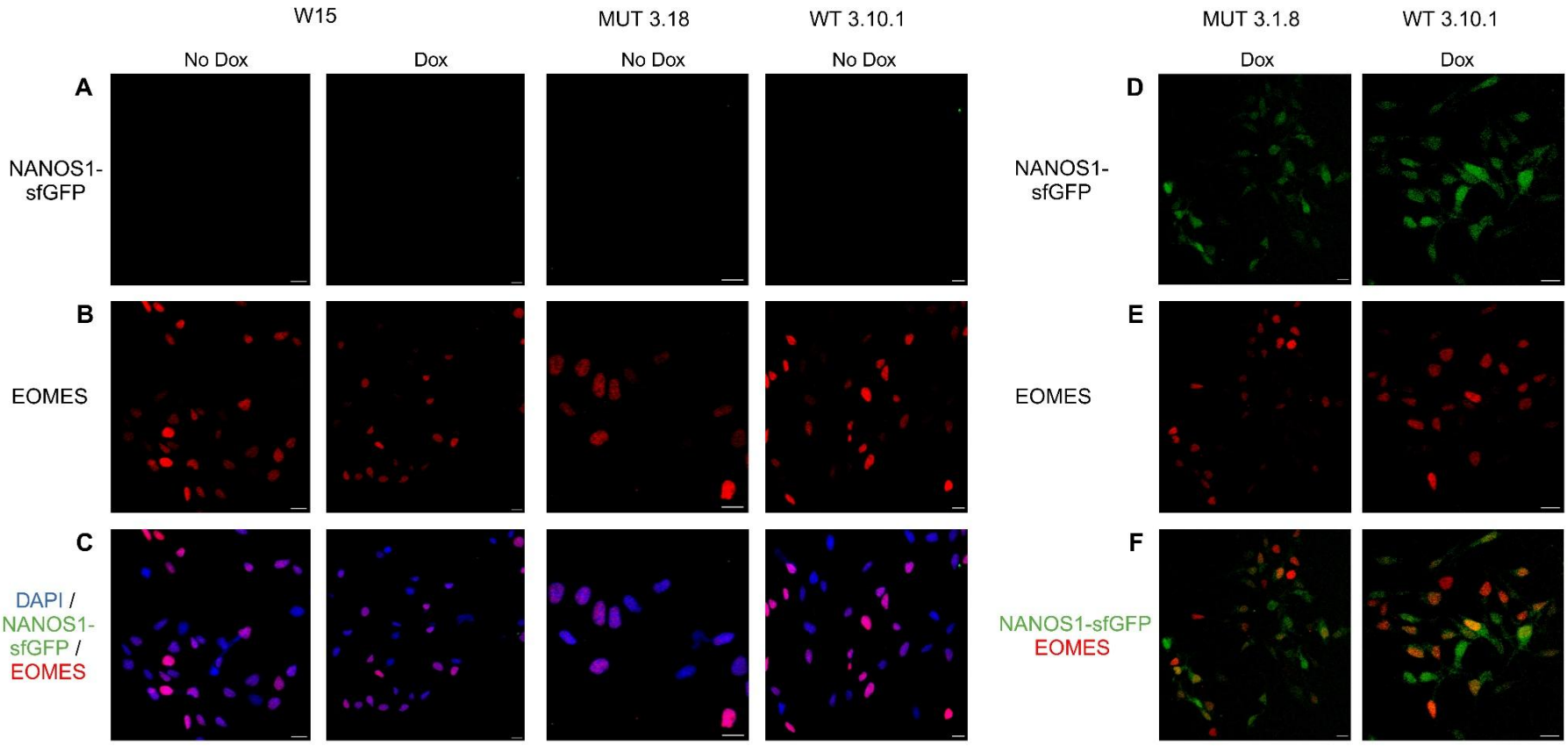

W15
MUT 3.18
WT 3.10.1
MUT 3.1.8
WT 3.10.1
No Dox
Dox
No Dox
No Dox
Dox
Dox
A
NANOS1-sfGFP
B
EOMES
C
DAPI / NANOS1-sfGFP / EOMES
D
NANOS1-sfGFP
E
EOMES
F
NANOS1-sfGFP
EOMES

Figure 22. Immunofluorescence analysis of the pre-me marker EOMES and NANOS1-sfGFP expression in the control cell line W15, MUT-NANOS1 3.1.8, and WT-NANOS1 3.10.1. A- NANOS1-sfGFP expression in pre-me cells fixed 6 h post-induction. This panel presents both uninduced and dox treated cell lines including control W15 cells to examine baseline NANOS1 expression during the early differentiation process. **B**- Immunostaining for EOMES in pre-me cells, focusing on cell lines before dox treatment and in both treated and untreated W15 control cells. **C**- Nuclei stained with DAPI. This panel overlays DAPI, NANOS1-sfGFP, and EOMES fluorescence, illustrating the co-localization of these markers in uninduced cell lines and across both treatment conditions in W15 control cells. **D**- NANOS1-sfGFP expression in dox-treated pre-me cells of MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 cell lines, fixed 6 h post-induction. The addition of dox at 500 ng/ml induces NANOS1 overexpression: Differential levels of NANOS1-sfGFP expression in MUT- and WT-NANOS1cell lines. **E**- Immunostaining for EOMES in dox-treated pre-me cells of both MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 to illustrate how EOMES expression is modulated following NANOS1 overexpression in both cell lines. **F**- Overlay of NANOS1-sfGFP and EOMES fluorescence in dox-treated cell lines, showing the co-expression and spatial overlap between these two markers. Scale bar: 70 µm (magnified for MUT-NANOS1 3.1.8 cell line without dox to provide enhanced visualization of marker expression).

Higher magnification was employed to closely investigate EOMES expression levels in the cell lines selected for high-throughput sequencing, specifically MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 (**Fig. 23A-C**). In contrast to the supplementary assessed cell lines MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1 (**Fig. 22D**), NANOS1-sfGFP expression levels were comparable between MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 cell lines, visible after dox induction (**Fig. 23A**), indicating that NANOS1 protein was similarly expressed in these cell lines upon induction. However, prior to dox treatment, the expression of EOMES was significantly higher in MUT-NANOS1 3.2.5.2 compared to WT-NANOS1 1.2.5.1 (**Fig. 23B-C**), further highlighting cell line variation in EOMES expression levels without the influence of protein overexpression.

Unlike WT-NANOS1 3.10.1 and MUT-NANOS1 3.1.8 cell lines, which were generated using different transposon-to-transposase ratios, both MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 cell lines were edited using an identical transposon-to-transposase ratio (2.5:1) (**Tab. 2**). Therefore, the differences in EOMES expression prior to dox treatment cannot be attributed to the ratio of integration and instead may reflect other factors such as cell line expansion dynamics or variability introduced during the genome editing process. These cell line effects highlight the variability that can arise even under similar editing conditions and their potential influence on baseline marker expression.

Following dox induction, NANOS1 overexpression had divergent effects on EOMES expression in the two cell lines. In MUT-NANOS1 3.2.5.2, EOMES levels significantly decreased in MUT-NANOS1-expressing pre-me cells (**Fig. 23B-C**), suggesting that MUT-NANOS1 protein might disrupt the regulatory processes required for maintaining PGC competence *via* EOMES. This reduction in EOMES expression supports the hypothesis that the p.[Pro34Thr; Ser78del] mutations might negatively impact critical regulatory networks at the pre-me stage, impairing PGC specification.

Conversely, in WT-NANOS1 1.2.5.1 cell line, EOMES expression remained relatively stable following the overexpression of WT-NANOS1, especially within the nucleus of induced pre-me cells (**Fig. 23B-C**). This stability suggests that WT-NANOS1 does not exert the same inhibitory effect on EOMES,

allowing cells to maintain the expression levels necessary for PGC specification. These findings indicate that while cell line effects on EOMES expression are evident prior to dox induction, the primary influence on post-induction EOMES levels is determined by the nature of the overexpressed NANOS1 protein. The MUT-NANOS1 double variant appears to have a stronger suppressive effect on EOMES expression, which may compromise the ability of pre-me cells to maintain PGC competence.

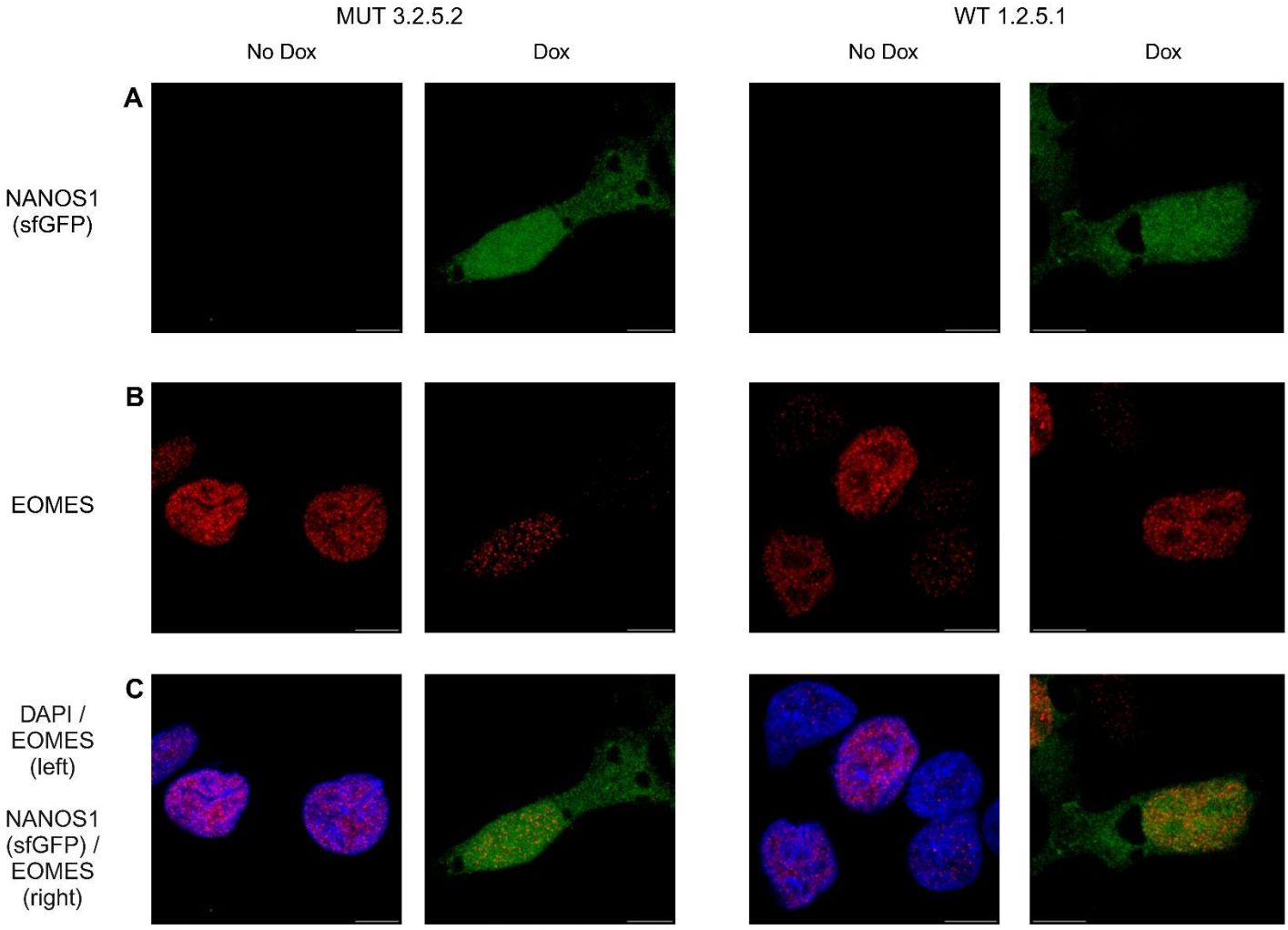


Figure 23. Immunofluorescence analysis of EOMES expression and NANOS1-sfGFP pre-me cells from MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 cell lines. A- NANOS1-sfGFP expression in pre-me cells, fixed 6 h post-induction. Dox was added at a concentration of 500 ng/ml at the beginning of pre-me induction to initiate NANOS1 overexpression. **B**- EOMES staining in pre-me cells, before and after dox treatment, to assess the impact of NANOS1 overexpression on EOMES expression in both MUT and WT cell lines. **C**- Nuclei stained with DAPI. This panel presents the overlap of DAPI, NANOS1-sfGFP, and EOMES fluorescence in uninduced cells, and the overlap of NANOS1-sfGFP and EOMES in their dox-treated counterparts, highlighting the spatial co-expression of these markers. Scale bar: 100 μm.

3.2.7 All selected cell lines (2 WT- and 2 MUT-NANOS1) express Day 4 EB markers SOX17, TFAP2C and PRDM14

Lastly, Day 4 EB visualization in both control and NANOS1 cell lines aimed at evaluating the potential for PGC specification before and after dox-induced NANOS1 overexpression. The primary marker for validation was SOX17, a critical TF for human PGC fate determination and one of the earliest markers of human PGC differentiation, distinguishing human germ cell specification from that of mice (Irie et al., 2015). The initial focus was on the control cell line W15 following dox induction, where both SOX17 and TFAP2C were examined as early markers of human PGCs (**Fig. 24A-H**). These markers have been shown to co-express with NANOS3, PRDM1, OCT4, and PRDM14, but notably, not with SOX2 at this stage of *in vitro* differentiation (Irie et al., 2015). In W15, the addition of dox did not alter the levels of NANOS3, SOX17, or TFAP2C, which were co-expressed in a subset of cells identified as hPGCLCs (**Fig. 24A-H**). This suggests that dox addition to Day 4 EBs does not significantly affect PGC marker expression in the control cell line W15.

In contrast, the cell lines WT-NANOS1 1.2.5.1 and MUT-NANOS1 3.2.5.2 showed distinct responses to NANOS1 overexpression. NANOS3-tdTomato expression was notably lower in both cell lines post-NANOS1 induction compared to the control W15 (**Fig. 24A and Fig. 24E**), suggesting that overexpression of NANOS1 in either form may interfere with NANOS3 expression levels in these clones. However, the SOX17 expression remained consistent across both cell lines (**Fig. 24G**), maintaining levels comparable to those observed in the control cell line W15 (**Fig. 24C**). This stable SOX17 expression suggests that the essential regulatory pathways governing PGC fate specification remain largely unaffected by NANOS1 overexpression in both WT and MUT cell lines. The cell lines also differed from the control W15 cell line in their expression of NANOS1-sfGFP. In the cell lines WT-NANOS1 1.2.5.1 and MUT-NANOS1 3.2.5.2, unlike in control W15, TFAP2C could not be simultaneously captured due to the sfGFP signal from NANOS1 overexpression occupying the available fluorophore range. Additionally, since NANOS3 was co-expressed with tdTomato, the fluorescence range was restricted, limiting detection to one marker at a time (**Fig. 24F** vs. **Fig. 24B**). In Day 4 EBs, NANOS1 fluorescence was detected in both the nuclear and cytoplasmic compartments. In the nucleus, NANOS1 co-localized with SOX17 and NANOS3, indicating a role in early PGC specification, while its cytoplasmic overexpression in somatic cells lacking these markers (**Fig. 24F**) likely reflects the non-specific nature of PiggyBac-mediated editing, which may have introduced NANOS1 into a wider range of cell types beyond those committed to germline fate. Overall, the MUT-NANOS1 3.2.5.2 cell line exhibited a homogeneous fluorescence pattern with uniform marker expression across the entire EB, whereas the WT-NANOS1 1.2.5.1 cell line showed greater variability in marker distribution, as observed by light sheet imaging (**Fig. 24F-G**). In MUT-NANOS1, SOX17 was uniformly distributed throughout the EB, unlike in WT-NANOS1, where it was confined to NANOS3-expressing human PGCs, mirroring the pattern seen with NANOS1-sfGFP (**Fig. 24F-G**). Since SOX17 is also an endoderm marker, its broader expression in the MUT-NANOS1 3.2.5.2 Day 4 EB suggests the presence of a potentially

substantial SOX17-positive, non-PGC population, implying that NANOS1 overexpression may drive differentiation beyond the intended PGC lineage.

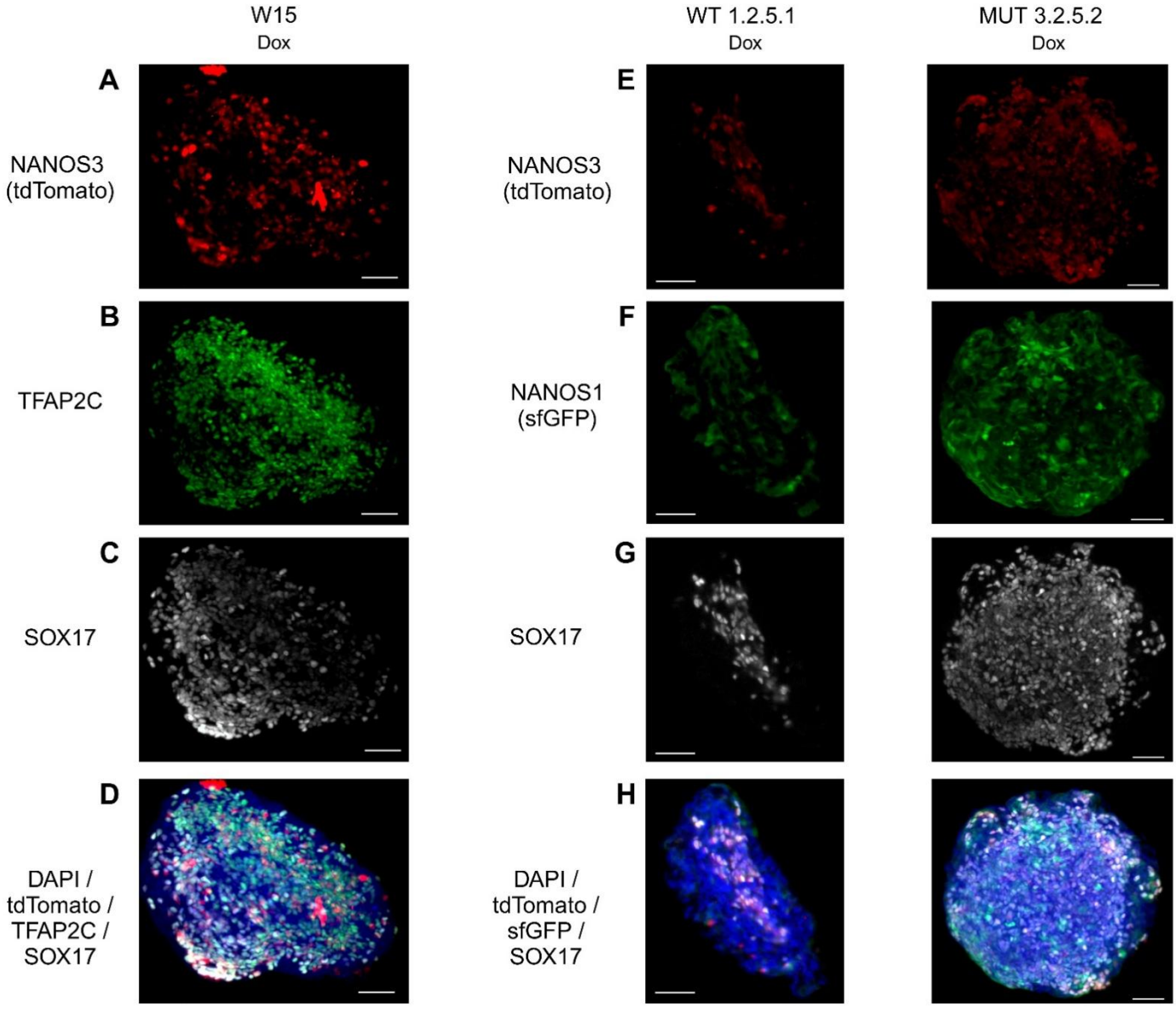


Figure 24. Light sheet imaging of Day 4 EBs post-dox treatment, assessing the expression of SOX17, TFAP2C, and NANOS3 (tdTomato) in the control cell line W15 and cell lines MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1. A- and **E**- NANOS3-tdTomato fluorescence, representing human PGCs within the EBs, was observed 24 h after dox addition, showing the distribution of NANOS3-positive cells in both control and PiggyBac cell lines. **B**- TFAP2C immunostaining in the control cell line W15, a key marker involved in human PGC development, showing its expression pattern in the presence of dox. **F**- NANOS1-sfGFP fluorescence, induced by dox in the cell lines MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1, visualizing the successful overexpression of NANOS1 and its localization within the EBs. **C**- and **G**- SOX17 staining across the control cell line W15 and both generated cell lines, MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1. This marker is essential for human PGC differentiation and was compared between the control and generated cell lines to assess any effects of NANOS1 overexpression on SOX17 expression. **D**- and **H**- Nuclei stained with DAPI. **D**- Overlap of DAPI, NANOS3 (tdTomato), TFAP2C, and SOX17 in the control cell line W15. **H**- Overlap of DAPI, NANOS3 (tdTomato), NANOS1-sfGFP, and SOX17 in the cell lines MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1. This allows for a comprehensive comparison of marker co-expression between the control and clonal cell lines. Scale bar: 100 µm.

The alternative cell lines, WT-NANOS1 3.10.1 and MUT-NANOS1 3.1.8, were also analysed post-dox induction for SOX17 expression to validate their suitability as supplementary cell lines for further analysis (**Fig. 25A-D**). Among both cell lines, NANOS3 fluorescence was notably lower compared to the control cell line W15 (**Fig. 24A**; **Fig. 25A**), which underscores the effect of the NANOS1 overexpression system on NANOS3 levels in the PiggyBac cell lines. This reduction in NANOS3 expression suggests that the introduction of NANOS1, particularly of the MUT-NANOS1, may be impacting the stability of PGC-specific marker expression. Interestingly, despite the expectation of lower NANOS1 expression in MUT-NANOS1 3.1.8 cell line due to the transposase-to-transposon ratio being 1:1 compared to the 1:10 ratio in WT-NANOS1 3.10.1 (**Tab. 2**), NANOS1 fluorescence was more pronounced in MUT-NANOS1 3.1.8 cell line (**Fig. 25B**). This heightened NANOS1 expression correlated with lower SOX17 fluorescence signal in MUT-NANOS1 3.1.8 EB (**Fig. 25C-D**). The elevated expression of MUT-NANOS1 in this cell line, contrary to expectations, suggests that the MUT-NANOS1 protein may be more stable or accumulate differently within the cell. This overexpression of MUT-NANOS1 likely disrupts the balance of TFs critical for proper PGC differentiation, which could explain the reduced SOX17 expression. The diminished SOX17 signal, a key marker of hPGCLCs, indicates that MUT-NANOS1 may be negatively impacting the PGC specification process, potentially pushing the cells toward an alternative fate or impairing their ability to differentiate correctly into PGC precursors. Moreover, despite these differences, both WT-NANOS1 3.10.1 and MUT-NANOS1 3.1.8 cell lines demonstrated an overlap between NANOS3 and SOX17 fluorescence, indicating that the core PGC markers were still present, albeit at reduced levels (**Fig. 25D**). The overlap between these two key PGC markers is critical, as it confirms the presence of cells committed to the PGC lineage. However, similarly as the cell lines in **Fig. 24**, NANOS1 fluorescence also overlaps with the somatic population (**Fig.25B**), indicating that the overexpressed protein extends into soma-like cell types. This dual expression suggests an unintended consequence of PiggyBac transposon-mediated editing, where non-specific integration and expression result in NANOS1 overexpression in non-target cell populations.

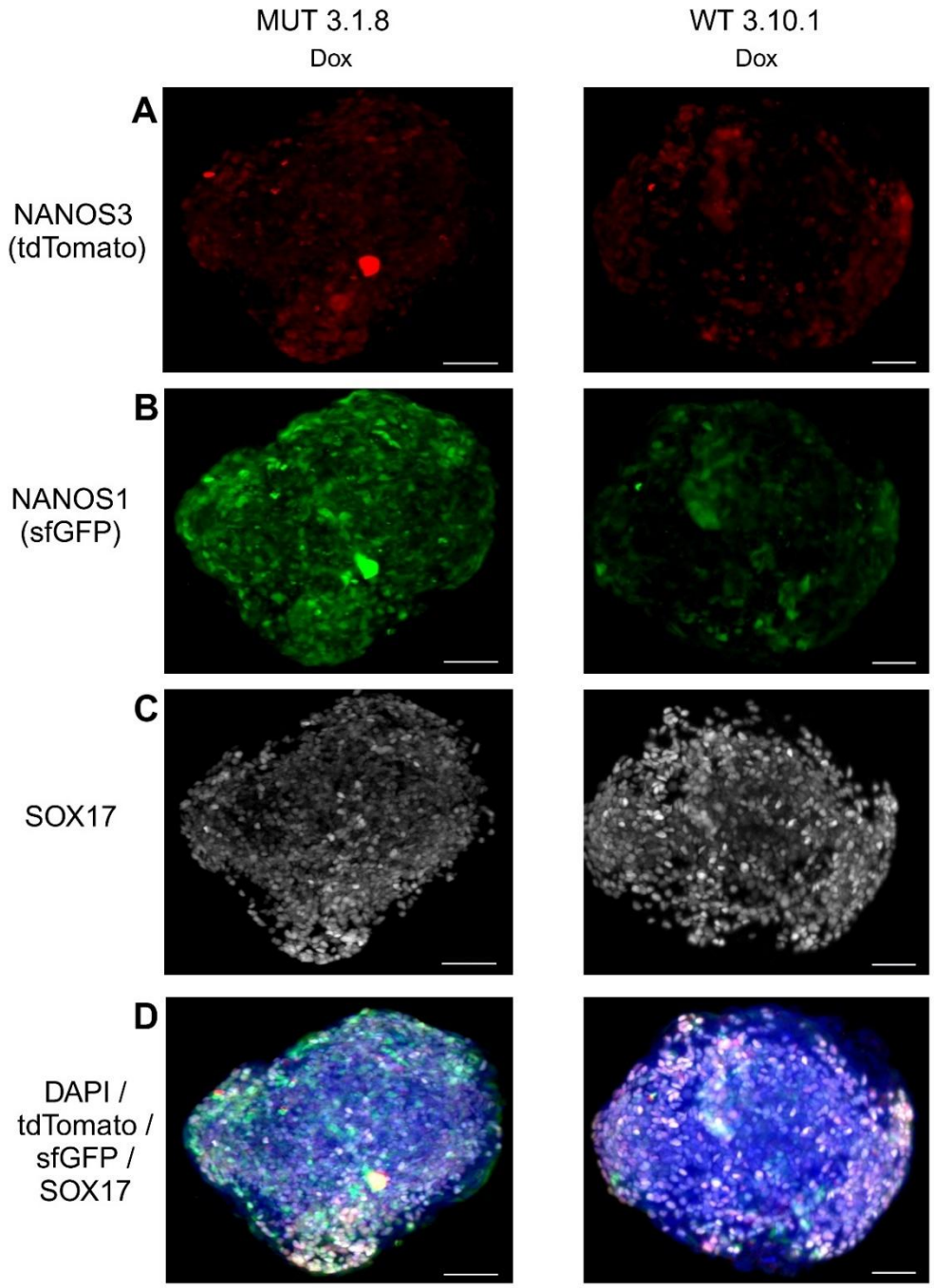


Figure 25. Light sheet imaging of Day 4 EBs post-dox treatment, focusing on the co-expression of SOX17, NANOS3 (tdTomato), and NANOS1-sfGFP in cell lines MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1. A- NANOS3-tdTomato fluorescence representing huma PGCs within the EBs, imaged 24 h after dox addition. The dox treated timepoints capture the response of NANOS3-positive cells to NANOS1 overexpression. **B**- NANOS1-sfGFP fluorescence, induced via dox treatment, visualizes the overexpression of NANOS1 in the cell lines MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1. Temporal overexpression of NANOS1 highlights the spatial distribution and intensity of NANOS1-sfGFP in EBs. **C**- SOX17 marker staining in EBs post-fixation and embedding in capillaries. The SOX17 marker in both cell lines post-dox is used to assess the specification of human PGCs in response to NANOS1 overexpression. **D**- Nuclei stained with DAPI. Overlay of DAPI, NANOS3 (tdTomato), NANOS1-sfGFP, and SOX17 fluorescence, highlighting the co-expression of these markers within the cell lines MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1. The overlay analysis reveals the spatial relationships between germline markers (NANOS3 and SOX17) and NANOS1 overexpression. Scale bar: 100 μm.

PRDM14, another key marker validated in Day 4 EBs, is critical for epigenetic reprogramming and maintaining pluripotency in human PGC-like cells; knockdown studies using a degron system revealed that loss of PRDM14 reduces PGC specification efficiency and disrupts the transcriptomic profile during hESC differentiation (Sybirna et al., 2020). Considering the importance of PRDM14 in human PGC differentiation, its expression was evaluated both before (**Fig. 26A-C** and **Fig. 28**) and after dox-induced NANOS1 overexpression (**Fig. 27A-G** and **Fig. 28**). Prior to dox induction, PRDM14 and NANOS3 were expressed at comparable levels in the control cell line W15 and the MUT-NANOS1 3.2.5.2 cell line, indicating equivalent baseline pluripotency and PGC competence (**Fig. 26A-B**). In contrast, the WT-NANOS1 1.2.5.1 cell line displayed relatively lower PRDM14 fluorescence (**Fig. 26A-B**), which aligns with earlier observations at the pre-me stage, suggesting that this cell line consistently exhibits lower pluripotency, which persists even on Day 4 of PGC specification. Notably, across all cell lines, PRDM14 and NANOS3 fluorescence overlapped within the nuclei of the PGCLCs within the EBs, confirming the nuclear localization and co-expression of these key markers (**Fig. 26C**). The nuclear co-localization of PRDM14 and NANOS3 is essential, as PRDM14 is involved in maintaining pluripotency and repressing somatic gene expression, while NANOS3 is a critical marker for PGC specification. Overall, the lower PRDM14 expression in the uninduced WT-NANOS1 1.2.5.1 cell line (**Fig. 26B**) implies a reduced capacity for epigenetic reprogramming, potentially leading to inefficient suppression of somatic differentiation pathways, thus limiting the cell line's overall PGC competence.

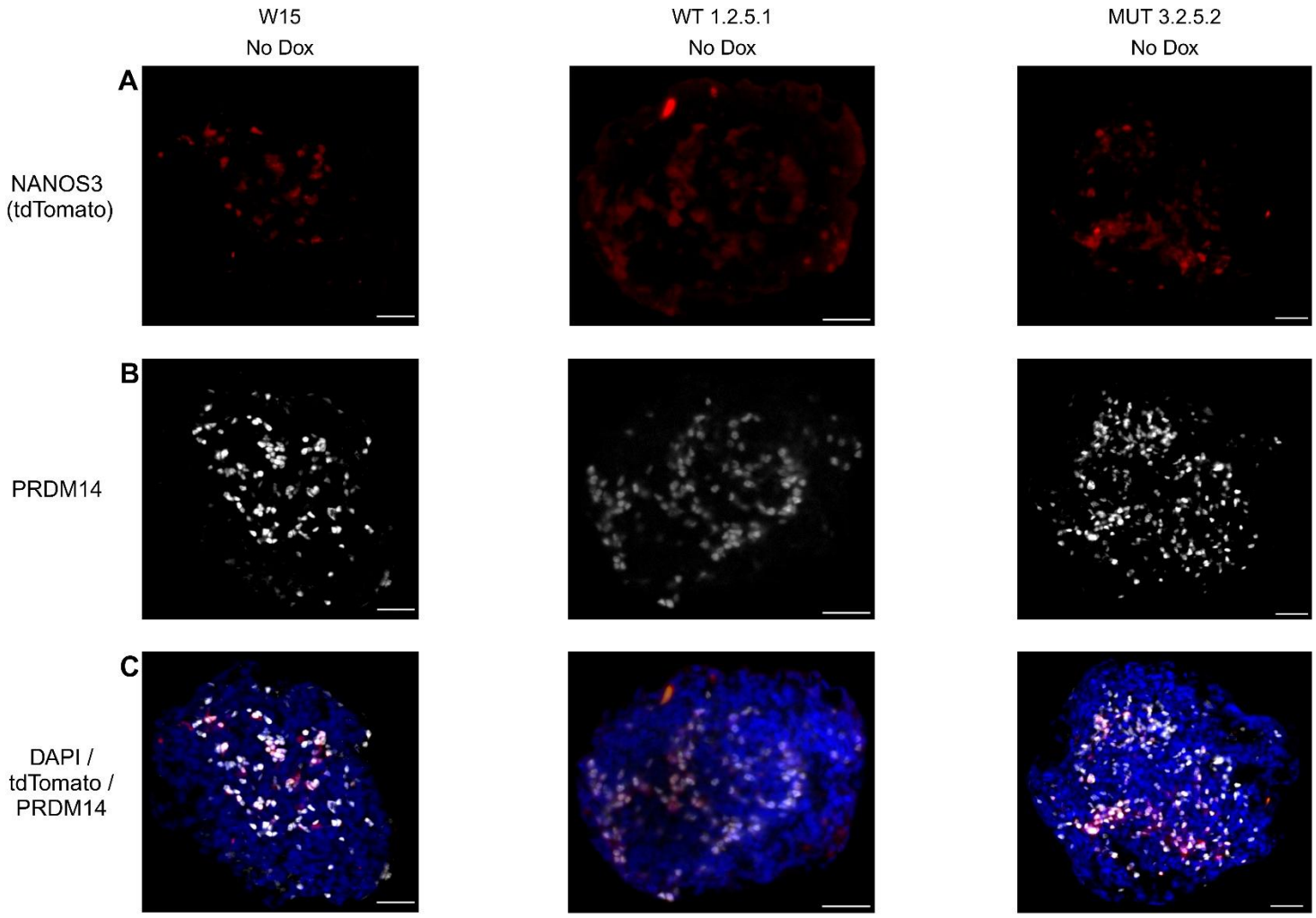


Figure 26. Light sheet imaging of Day 4 EBs before dox treatment, highlighting PRDM14 staining in relation to NANOS3 (tdTomato) expression. This was carried out in the control cell line W15 and PiggyBac cell lines MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1 in the absence of NANOS1 protein induction (no dox). **A**- NANOS3-tdTomato fluorescence representing human PGCs within the EBs after 4 days of differentiation in the presence of specific cytokines. This fluorescence identifies the human PGC population and allows for spatial localization within the EBs. **B**- PRDM14 marker staining within EBs post-fixation and embedding in gel capillaries. PRDM14 is a key regulator of epigenetic reprogramming and pluripotency in PGCs, and its expression is analyzed to assess the germline specification potential in both control and generated cell lines without NANOS1 overexpression. **C**- Nuclei stained with DAPI for cellular structure visualization. The panel shows the overlap of DAPI, NANOS3 (tdTomato), and PRDM14 fluorescence, demonstrating the co-expression and spatial relationships of these key PGC markers in the EBs. This co-localization provides insight into the coordination of PGC marker expression in the absence of NANOS1 induction. Scale bar: 100 μm.

Following the addition of dox (**Fig. 27A-G**), the control cell line W15 exhibited no significant changes in the levels of PRDM14 (**Fig. 27B**). Both PRDM14 and NANOS3 maintained expression levels comparable to those observed prior to dox treatment (**Fig. 26A-C**). Furthermore, the markers overlap within the nuclei of human PGCs in the EB, where they consistently remain localized, as expected for cells progressing through normal PGC differentiation (**Fig. 27C**). This stable expression pattern suggests that the dox induction itself does not influence the baseline expression of these key PGC markers in the control cell line.

In contrast, the MUT 3.2.5.2 cell line exhibited a marked reduction in both NANOS3 and PRDM14 expression following MUT-NANOS1 overexpression (**Fig. 26A-B**, **Fig. 27D** and **Fig. 27F).** Prior to dox treatment, MUT-NANOS1 3.2.5.2 cell line displayed pluripotency levels comparable to the control cell line, as indicated by similar expression levels of PRDM14 and NANOS3 (**Fig. 26A-C**). However, the observed decline in these markers post-induction points to a specific impact of MUT-NANOS1 overexpression on PGC specification pathways. The reduction in PRDM14 and NANOS3 suggests that the MUT-NANOS1 may disrupt the transcriptional regulation required for maintaining pluripotency and promoting proper PGC differentiation as PRDM14 is crucial for repressing somatic signals and supporting the epigenetic reprogramming required for PGC fate (Irie et al., 2015).

For the WT 1.2.5.1 cell line, the effects of dox-induced overexpression were more subtle. NANOS3 and PRDM14 levels remained relatively stable, with only a slight improvement in fluorescence intensity following WT-NANOS1 overexpression when compared to the MUT-NANOS1 cell line (**Fig. 26A-B**; **Fig. 27D** and **Fig. 27F**). Although there was a minor increase in these PGC markers, their expression levels were still qualitatively lower than those observed in the control cell line W15. This persistent lower expression suggests that the cell line effect in WT 1.2.5.1, identified at earlier differentiation stages, continues to influence the cells' ability to fully engage the transcriptional program required for PGC specification, even after WT-NANOS1 overexpression. The modest increase in PRDM14 and NANOS3 expression suggests that, although WT-NANOS1 exerts a positive effect, it is insufficient to overcome the cell line's intrinsic limitations in supporting robust human PGC differentiation.

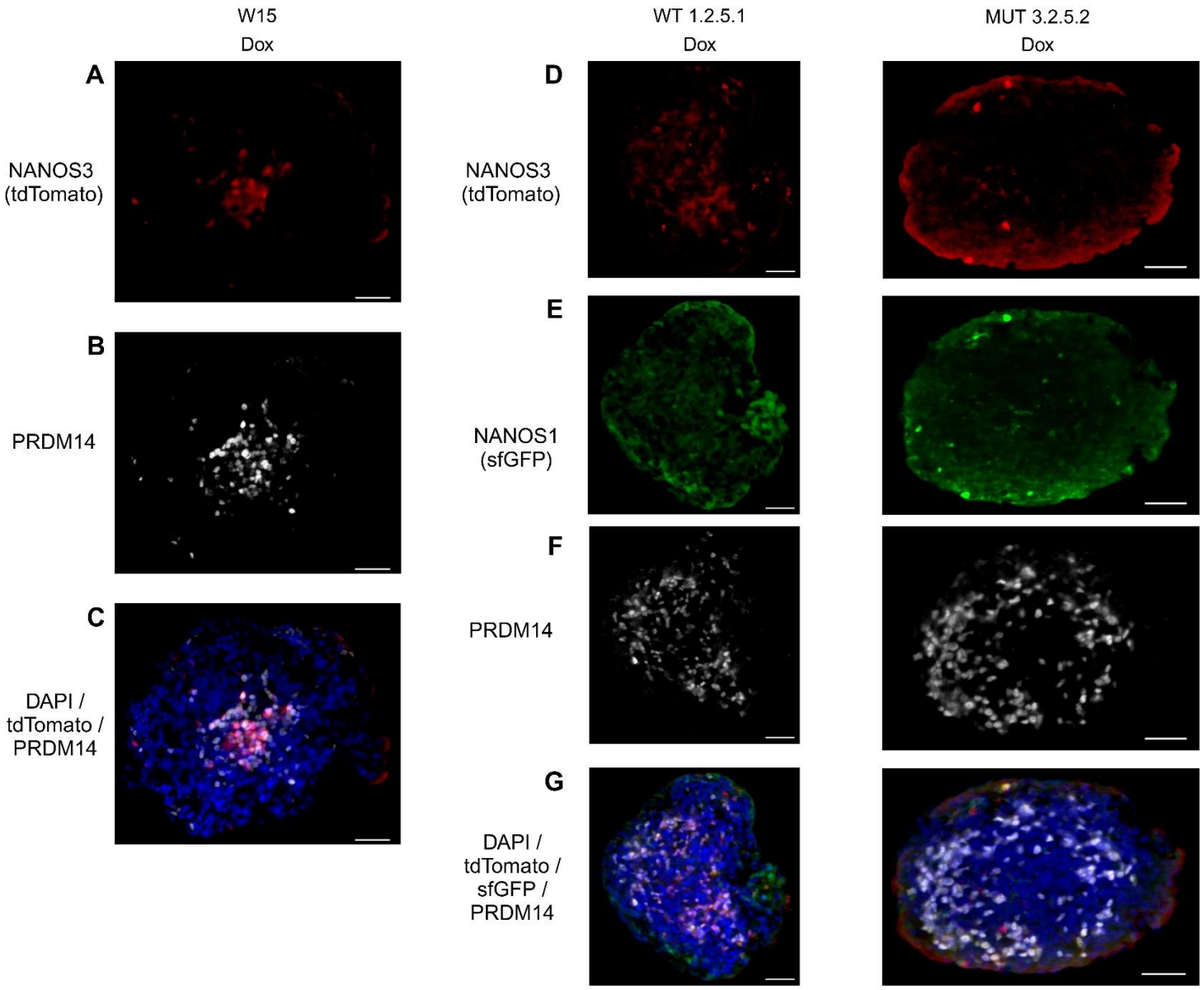


Figure 27. Light sheet imaging of Day 4 EBs post-dox treatment, highlighting PRDM14 staining and its co-expression with NANOS3 (tdTomato) and NANOS1-sfGFP overexpression. This assay was carried on dox-treated control cell line W15 and cell lines MUT-NANOS1 3.2.5.2 and WT-NANOS1 1.2.5.1. **A**- and **D**- NANOS3-tdTomato fluorescence representing human PGCs within the EB following 4 days of differentiation in the presence of specific cytokines. Dox was added at a concentration of 500 ng/ml 24 h prior to EB collection, fixation, and staining. **B**- PRDM14 marker staining in the dox-treated control cell line W15, following fixation and embedding in gel capillaries to assess its expression in the context of human PGC differentiation. **C**- Overlap of DAPI (nuclear stain), NANOS3 (tdTomato), and PRDM14 fluorescence in dox-treated W15 EB, illustrating the co-localization of these markers in control W15 cells. **E**- NANOS1-sfGFP fluorescence induced by dox through the PiggyBac and Tet-ON cassette systems, showing successful overexpression of NANOS1 in the cell lines. **F**- PRDM14 staining in dox-treated MUT 3.2.5.2 and WT 1.2.5.1 cell lines, showing the effects of NANOS1 overexpression on PRDM14 expression in the respective EBs. **G**- Nuclei stained with DAPI. This panel shows the overlap of DAPI, NANOS3 (tdTomato), NANOS1-sfGFP, and PRDM14 fluorescence in dox-treated MUT 3.2.5.2 and WT 1.2.5.1) cell line, providing insight into how NANOS1 overexpression influences the expression of key PGC and pluripotency marker PRDM14 during human PGC differentiation. Scale bar: 100 μm.

In supplementary cell lines, similar PRDM14 expression patterns were observed (**Fig. 28A-D**). Prior to dox induction, MUT-NANOS1 3.1.8 exhibited higher PRDM14 levels than WT-NANOS1 3.10.1, with its expression comparable to that of the control cell line W15, indicating similar baseline pluripotency (**Fig. 26B, Fig. 28A** and **Fig. 28C**). This suggests that prior to dox induction, MUT cell lines retain their pluripotency and PGC potential comparable to the baseline state of the control cell line W15.

However, following dox induction and subsequent overexpression of the MUT-NANOS1 protein, a marked reduction in PRDM14 levels was observed in MUT 3.1.8, mirroring the trend seen in MUT 3.2.5.2 cell line (**Fig. 27F** and **Fig. 28A**). This decline in pluripotency, as indicated by reduced PRDM14 expression, suggests that MUT-NANOS1 overexpression may exert a negative effect on the regulatory pathways necessary for maintaining pluripotency and promoting human PGC differentiation. Interestingly, this decrease in pluripotency is consistent between both independently generated MUT cell lines, despite differences in the transposon-to-transposase ratios used during their editing (1:1 in MUT 3.1.8 versus 2.5:1 in MUT 3.2.5.2 **Tab. 2**). This consistency points to a robust effect of the MUT-NANOS1 protein in disrupting PGC specification processes, regardless of the variation in transposon integration ratios, underscoring its significant role in modulating pluripotency.

In contrast, the WT cell lines, including WT 3.10.1, displayed lower PRDM14 levels both before and after dox induction when compared to the control cell line W15 and the MUT cell lines (**Fig. 28C**, **Fig. 27F** and **Fig. 26B**). This finding is consistent with the behavior of WT 1.2.5.1 cell line, where PRDM14 expression was also reduced relative to the control (**Fig. 27F** and **Fig. 26B**). In WT 3.10.1, PRDM14 expression did not significantly change following WT-NANOS1 overexpression, suggesting that the reduced pluripotency observed in the WT-NANOS1 cell lines is primarily due to intrinsic clonal effects rather than the overexpression itself. This indicates that the baseline pluripotency of the WT cell lines, reflected by PRDM14 as well as other markers in pre-me and stem cells stages, may be inherently lower, limiting their capacity for effective hPGCLC differentiation regardless of NANOS1 induction. Moreover, the persistent low PRDM14 expression implies that genomic or epigenetic factors in these WT clones may restrict the full activation of pluripotency and germline differentiation pathways. Although WT-NANOS1 exhibited a slight positive effect on PRDM14 expression in WT 1.2.5.1 (**Fig. 27F**), the overall lack of a substantial increase suggests that cell line characteristics exert a stronger influence on differentiation potential than WT-NANOS1 overexpression.

Overall, the analysis of PRDM14 expression as well as SOX17 in both MUT and WT cell lines highlights significant differences in their response to NANOS1 overexpression. The MUT cell lines show a clear decline in pluripotency and pre-me markers following MUT-NANOS1 induction, whereas the WT cell lines exhibit persistently lower PRDM14 and SOX17 levels that are more reflective of cell line limitations rather than an effect of NANOS1 expression. These findings emphasize the importance of cell line variability and the nature of NANOS1 in determining the efficiency of PGC specification and pluripotency maintenance, with the MUT- NANOS1 protein appearing to have a more disruptive role in this process than its WT counterpart, as highlighted by immunofluorescence validation.

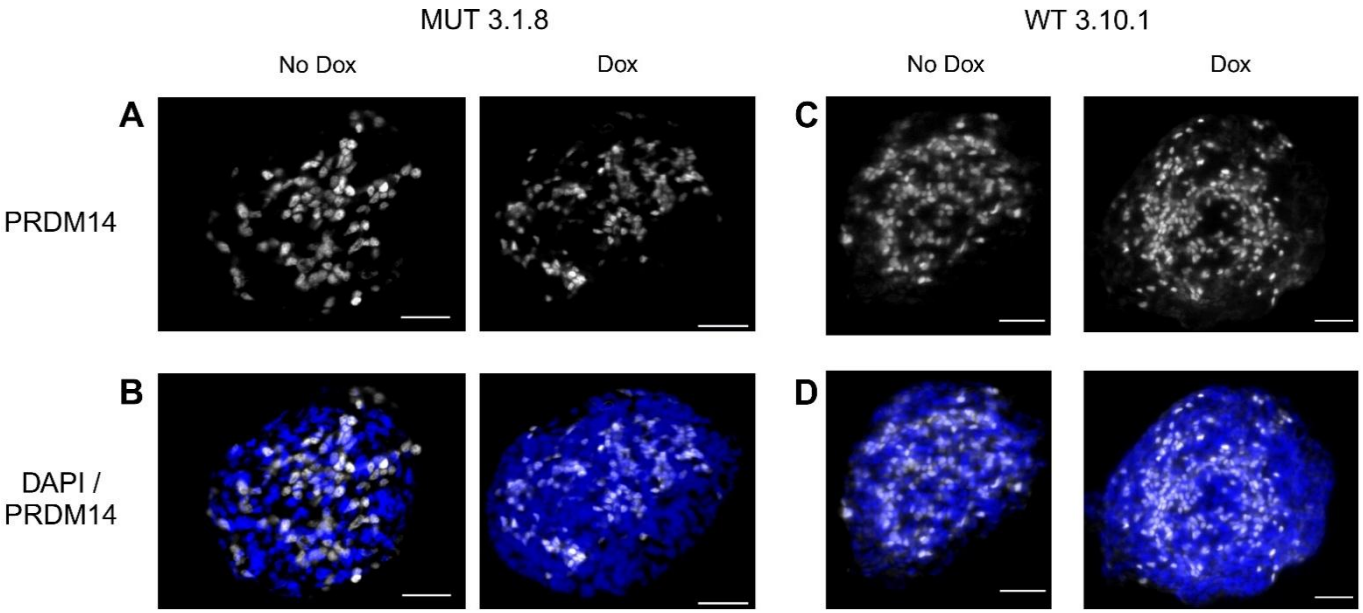


Figure 28. Light sheet imaging of Day 4 EBs highlighting PRDM14 expression in supplementary cell lines MUT-NANOS1 3.1.8 and WT-NANOS1 3.10.1. A- and **C**- PRDM14 marker staining in uninduced and dox-treated EBs from the cell lines MUT 3.1.8 **A**- and WT 3.10.1 **C**-. Dox was added at a concentration of 500 ng/ml 24 h prior to fixation, staining, and embedding in gel capillaries to induce NANOS1 overexpression. This panel illustrates the effects of dox-based NANOS1 induction based on PRDM14 expression, a key regulator of germline specification and epigenetic reprogramming. **B**- and **D**- Nuclei stained with DAPI. This panel shows the overlap of DAPI and PRDM14 fluorescence in MUT 3.1.8 **B**- and WT 3.10.1 **D**- cell lines, both before and after dox-induced NANOS1 protein overexpression. Scale bar: 100 μm.

4. MUT-NANOS1 binds significantly more mRNAs than the WT protein, demonstrating a gain-of-function via an expanded 3′UTR RNA interactome as revealed by eCLIP

In an effort to investigate the bound RNA interactome of both WT-NANOS1 and MUT-NANOS1 proteins, the eCLIP assay was employed. This assay was conducted exclusively at the pre-me stage using the respective cell lines: WT 1.2.5.1 and MUT 3.2.5.2. This developmental stage was selected because the stringent washing steps required for immunoprecipitation necessitate a substantial amount of starting material, approximately 20 mln cells, which is achievable only at the pre-me stage.

4.1 Validation of enhanced crosslinking immunoprecipitation (eCLIP) procedure by using the WT-NANOS1 1.2.5.1 cell line in pre-me stage and anti-FLAG antibody

To validate the capture of the RNA interactome, an anti-FLAG antibody was utilized on the WT 1.2.5.1 cell line (**Fig. 29A–C**). As shown in **Fig. 29A-B**, the NANOS1-3×FLAG fusion protein was successfully captured *via* western blot analysis (10% SDS-PAGE gel) only when the WT-NANOS1 expression was induced with dox, regardless of whether the cells were cross-linked (XL) or non-cross-linked (noXL). No protein bands were observed under the same conditions in the uninduced cell line (**Fig. 29A**), confirming the dox-dependent expression and capture of the protein.

Further analysis involved visualizing bound RNA on a nitrocellulose blot using biotin-based imaging following RNAse I digestion of the extracted RNA (**Fig. 29C**). As expected, only the XL sample produced an RNA smear, indicating successful precipitation of the protein bound RNA, despite the digestion being slightly incomplete as evidenced by the patchy appearance of the smear. The non-XL sample did not exhibit an RNA smear, even though protein precipitation was confirmed (**Fig. 29A–B**), suggesting that RNA-protein interactions were preserved only under XL conditions (**Fig. 29C**).

A lighter RNA smear from a previous positive control run (XL) further supported the reproducibility of RNA interactome extraction post-precipitation, with the reduced intensity correlating to a lower amount of starting digested RNA (less than 20 µg) (**Fig. 29C**). Additionally, the uninduced sample (WT IP uninduced) did not generate an RNA smear under either XL or non-XL conditions, reinforcing the stringent control of the Tet-ON system over NANOS1 overexpression (**Fig. 29C**).

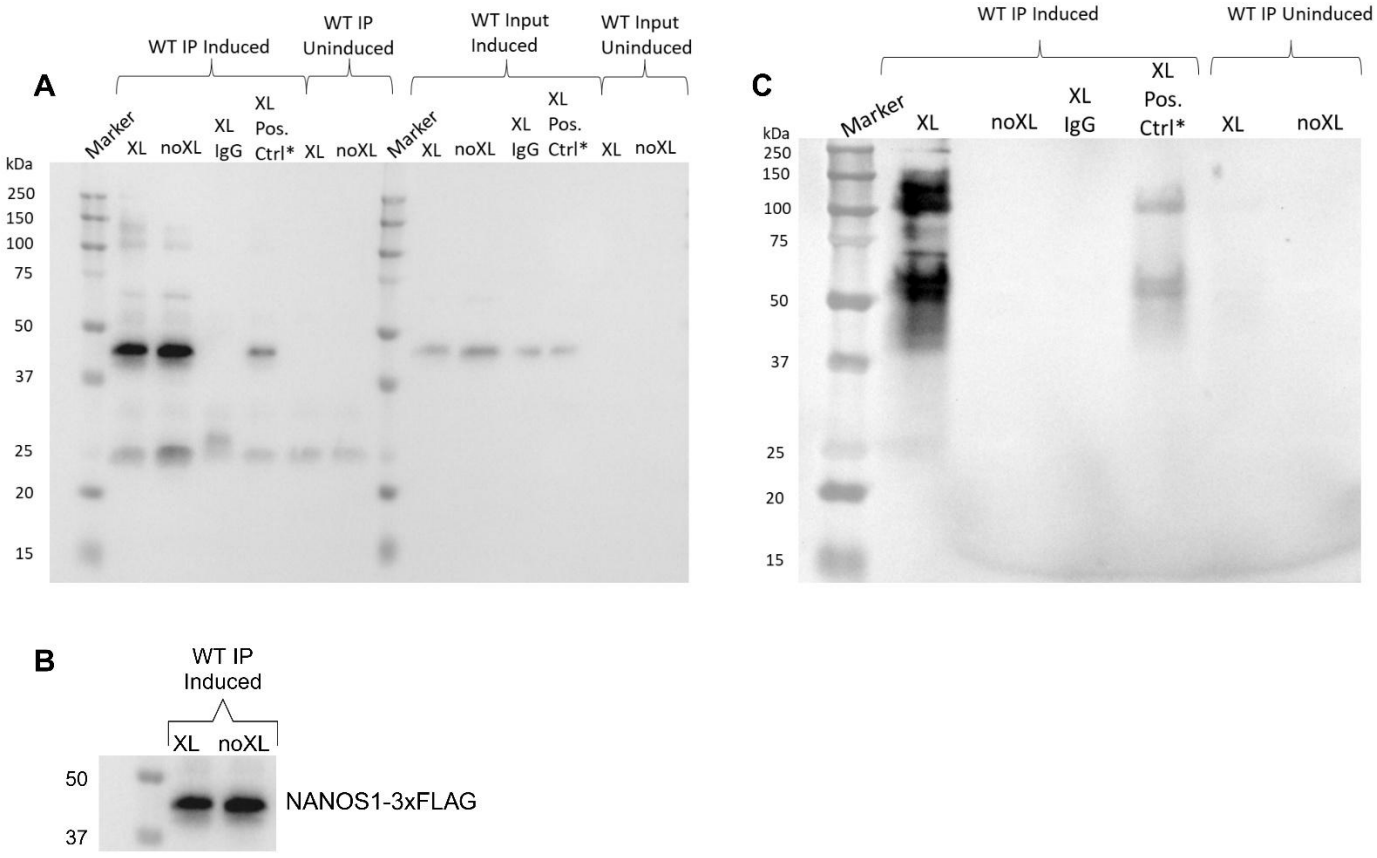


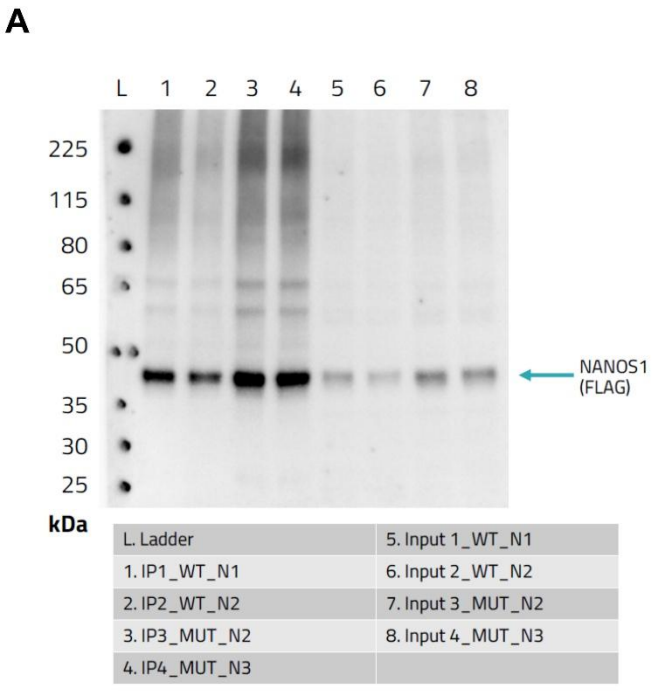
Figure 29. Validation of eCLIP protocol: Immunoprecipitation (IP) of overexpressed NANOS1 protein *via* anti-FLAG antibody. A- Western blot analysis of immunoprecipitated WT-NANOS1 cell line overexpressing NANOS1 conjugated with a 3×FLAG tag. An anti-FLAG® M2 antibody (F1804, Sigma) diluted 1:1500 was employed. Protein fractionation was performed using a 10% SDS-PAGE gel, followed by transfer to a PVDF membrane. Both IP and input samples were loaded. **B**- Enlarged view of the NANOS1-3×FLAG band from both XL (at 400 mJ/cm²) and noXL lysates of WT IP induced with 500 ng/ml dox. **C**- Visualization of IP samples dox induced and uninduced conditions using biotin-based chemiluminescence. Samples were similarly run on a 10% SDS-PAGE gel and transferred to a nitrocellulose membrane. 20 µg of extracted total RNA was digested with RNase I at a ratio of 4 µg RNaseI per 5 µg RNA. Note: Positive Control, WT 1.2.5.1 XL dox induced lysate from a previous experiment with promising results.

4.2 NANOS1-bound RNA identification in WT-1.2.5.1 and MUT-3.2.5.2 cell lines using eCLIP (adapted and performed from EclipseBio)

The eCLIP protocol described by EclipseBio (see **Methods**, **Fig. S14**) was applied to two WT-NANOS1 replicates (WT_N1 and WT_N2) and two MUT replicates (MUT_N2 and MUT_N3). Using the validated anti-FLAG antibody (**Fig. 29A–C**), the 3×FLAG-tagged NANOS1 protein was successfully precipitated in all replicates (**Fig. 30A**), consistently appearing at an approximate molecular weight of 40 kDa.

Visualization of the bound RNA on nitrocellulose membranes (**Fig. 30B**) confirmed its presence in all replicates for both WT- and MUT-NANOS1. However, variability was observed between the two WT replicates (lanes 1 and 2 in **Fig. 30B**); specifically, lane 1 exhibited a slightly fainter RNA smear compared to lane 2. This discrepancy suggests potential differences in RNA cross-linking efficiency or technical factors affecting these replicates.

In contrast, the MUT replicates (lanes 3 and 4 in **Fig. 30B)** demonstrated greater consistency. Lane 3 (IP3_MUT_N2) displayed a strong, continuous RNA smear with high signal intensity, indicative of effective RNA–protein cross-linking or a potentially more enriched RNA interactome. Similarly, lane 4 (IP4_MUT_N3) showed a robust RNA smear comparable in intensity to lane 3. Therefore, the two MUT replicates were more reproducible, exhibiting strong and uniform smears, which indicates consistency in RNA–protein interactions or cross-linking efficiency between these samples.



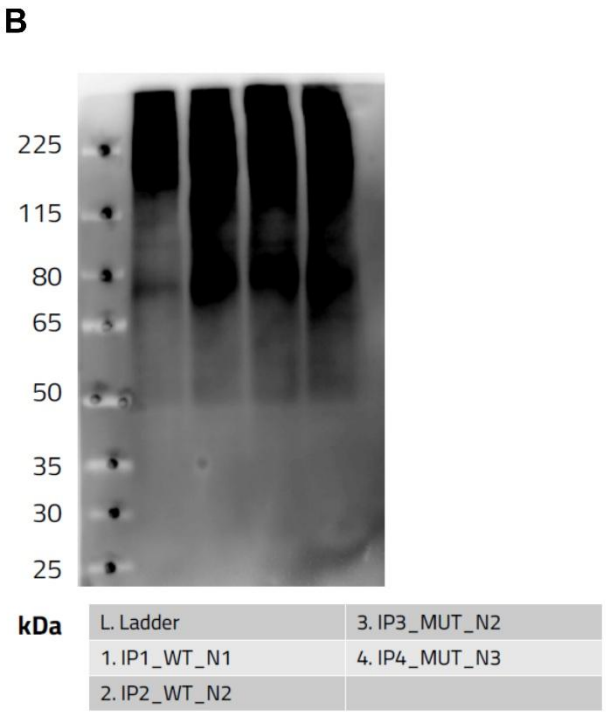


Figure 30. Immunoprecipitation and RNA visualization of NANOS1-3×FLAG bound RNAs from two IP_WT and two IP_MUT pre-me replicates. A- Western blot analysis of immunoprecipitated NANOS1-3×FLAG from eCLIP pre-me cells. 10% of IP samples (samples 1–4) and 1% of input samples (samples 1–4) were run on NuPAGE 4–12% Bis-Tris protein gels, transferred to a PVDF membrane, and probed with a 1:3000 dilution of monoclonal ANTI-FLAG® M2 antibody and a 1:4000 dilution of TrueBlot anti-Mouse IgG (HRP). During eCLIP, only the protein region from approximately 40 kDa to 115 kDa (proteins up to 75 kDa in size) was isolated. **B**- RNA visualization of immunoprecipitated NANOS1-3×FLAG from eCLIP pre-me cells. 10% of IP samples were run on NuPAGE 4–12% Bis-Tris protein gels, transferred to a nitrocellulose membrane, and visualized via chemiluminescence for 1 minute. WT replicates are labelled as WT_N1 and WT_N2; MUT replicates are labelled as MUT_N2 and MUT_N3.

4.3 Differential RNA interactomes: WT-NANOS1 predominantly targets introns, whereas its MUT counterpart engages a broader RNA interactome that includes introns, 3′UTRs, and CDS

After library preparation and sequencing of RNA extracted from the western blot (**Fig. 30A**), site enrichment analysis of bound RNAs was performed for both WT- and MUT- replicates (**Fig. 31** and **Fig. 32**). Specifically, **Figure 31** (left panels, top and bottom) illustrates the number of enriched RNAs potentially NANOS1-bound sites across various genomic features. Both replicates (WT_IP_1 and WT_IP_2) exhibit strong enrichment in intronic regions, which contain the highest number of enriched binding sites. Specifically, NANOS1 WT_IP_1 displays over 3,000 enriched intronic sites, whereas NANOS1 WT_IP_2 has approximately 12,000. This disparity may reflect the quantitative differences observed in each replicate's RNA smear (**Fig. 30A**). Additionally, both replicates (**Fig. 31**, left top and bottom panels) exhibit moderate to small enrichment in coding sequence (CDS) regions, with over 2,500 enriched sites for replicate 2 and 1,000 enriched sites for replicate 1, indicating that NANOS1 also binds to the CDS sites of mRNAs. This suggests a potential role in regulating mRNA translation efficiency or stability. The 3' untranslated region (3'UTR) is another notable binding site where WT-NANOS1 exhibits moderate binding (**Fig. 31**, left top and bottom panels). Overall, 3'UTRs regulate mRNA degradation, stabilization and localization, so NANOS1 binding here suggests it modulates these processes similarly to other proteins (Richter and Sonenberg, 2005). Moderate enrichment (~1,000 sites per replicate) indicates WT-NANOS1 contributes to this regulation, though less extensively.

Other genomic regions show negligible enrichment, ruling out additional regulatory mechanisms for WT-NANOS1. The background panel (**Fig. 31**, top right panel) indicates that introns, encoding factors like miRNAs, long noncoding RNAs and harboring alternative promoter elements, are the most abundant feature, with far fewer 3′UTRs and CDS. Overall, WT-NANOS1's highly enriched intronic binding suggests a potentially prominent nuclear role, while its moderate binding to 3′UTRs and CDS, along with its known translational repressor activity, points to a secondary cytoplasmic role in mRNA regulation.

Lastly, regarding enrichment overlap between replicates (**Fig. 31**, bottom right panel), the 26.7-fold odds ratio reflects the high reproducibility of NANOS1's RNA-binding preferences across experiments. The variation in intensity of the RNA smears observed in the RNA visualization blot (**Fig. 30B**) corresponds to differences in the amount of pulled-down material, with WT_IP1 exhibiting lower RNA yield compared to WT_IP2. This is reflected in the lower number of enriched sites in the first replicate relative to the second

(**Fig. 31**, left top and bottom panels). Despite this variation, the sequencing data from the replicates demonstrate consistent RNA-binding preferences of NANOS1.

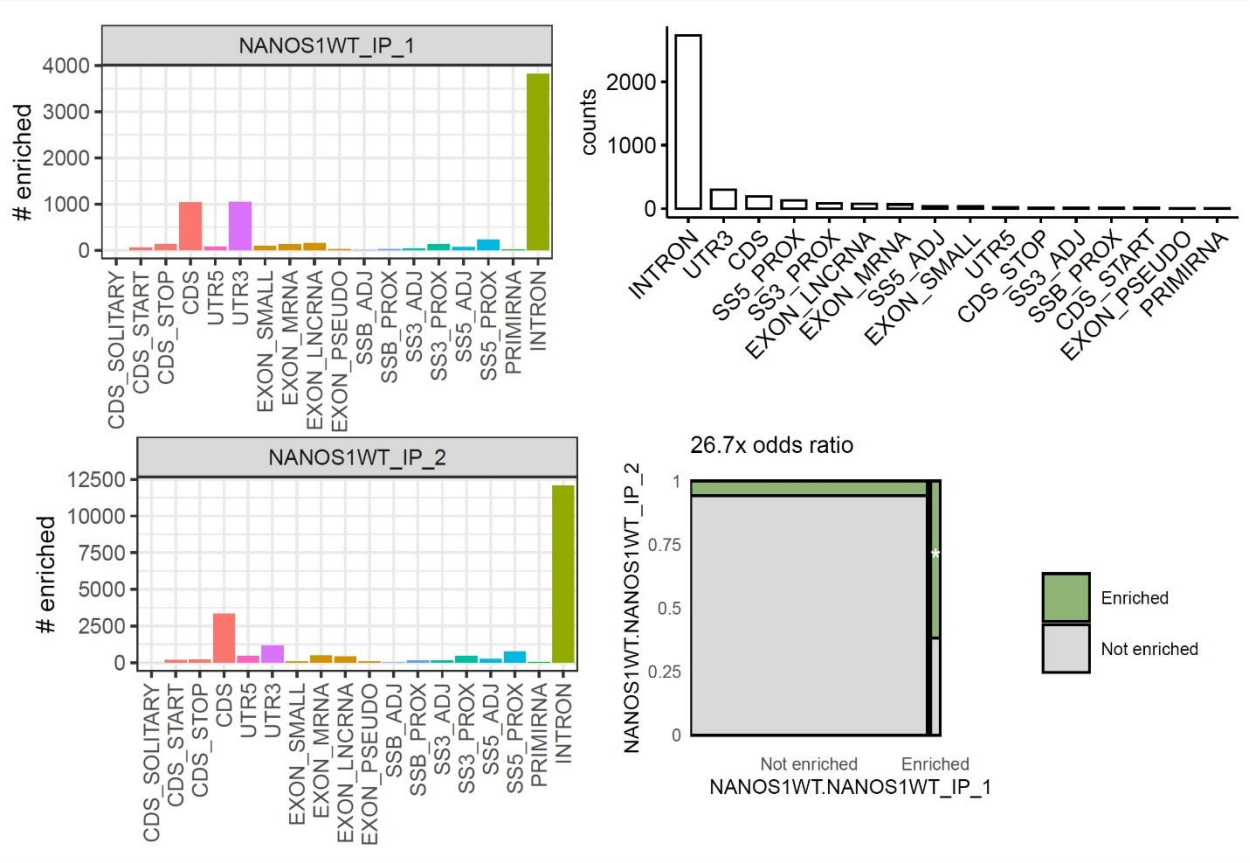


Figure 31. Enrichment of WT-NANOS1 binding across RNA features from two independent pre-me eCLIP replicates. The target RNA enrichment sites are highlighted from two independent immunoprecipitation (IP) experiments: NANOS1 WT_IP_1 (top left panel) and NANOS1 WT_IP_2 (bottom left panel). RNA regions are categorized by feature types, including CDS, UTRs, exons, introns, small RNAs, pseudogenes, lncRNAs, and splice sites (SS). The counts of enriched RNA features from both replicates are displayed in the top right panel, demonstrating that introns dominate the binding sites. The odds ratio plot (bottom right panel) compares the reproducibility of binding specificity across the RNA features between the two replicates. In this plot, the green region represents RNA features that are significantly enriched in both replicates (* denotes significance).

In contrast to the WT replicates, the enrichment analysis for the MUT replicates reveals a different pattern (**Fig. 32**). While intronic enrichment remains prominent, there is a higher significant enrichment in the 3'UTR and CDS compared to the WT samples. In MUT replicate 1 (MUT_IP_1; **Fig. 32**, top left panel), intronic binding is the most prominent with over 6,000 enriched sites. Although this represents a reduction compared to the WT counterpart, specifically replicate 1 (**Fig. 31**, top left panel), the MUT-NANOS1 protein retains a substantial ability to bind intronic regions.

Notably, 3'UTR is the next most enriched region and exhibits higher levels than in both WT replicates, with over 3,000 enriched sites surpassing CDS, which is the third most enriched region for MUT replicate 1. This shift implies an increased association of MUT-NANOS1 with functionalities related to mRNA stability, localization, and repression. In MUT replicate 2 (MUT_IP_2; **Fig. 32**, bottom left panel),

although intronic regions remain relatively high (over 3,000 enriched sites), 3'UTR enrichment is dominant and most abundant, with approximately 4,000 enriched sites, followed by CDS, which also shows higher enrichment compared to MUT replicate 1.

These observations, based solely on the eCLIP data, may reflect a functional shift in the MUT-NANOS1 protein toward 3'UTR mediated regulation. This is further supported by the background distribution of RNA features (**Fig. 32**, top right panel), where intronic regions and 3'UTR are the most abundant, followed by CDS, reflecting the overall composition of the RNA pool captured by both replicates. In pre-me cells primed for PGC development, MUT-NANOS1 preferentially binds introns, 3′UTRs, and CDS. The stronger enrichment at 3′UTRs and CDS suggests that post-transcriptional and translational regulation is maintained, potentially affecting the transition of MUT-NANOS1 pre-me cells into PGCs by altering the expression of essential mRNAs.

Lastly, the enrichment plot in **Fig. 32** (bottom right panel) highlights the overlap of enriched binding sites between the two MUT replicates, showing an 11.5-fold odds ratio, which is lower than that observed for WT replicates (27.6-fold odds ratio). This indicates that not all RNA-binding events are as consistent across both MUT replicates compared to the WT counterparts, potentially due to altered RNA-binding specificity, weaker or more unstable interactions, or variability in the subcellular localization of the MUT-NANOS1 protein. Nonetheless, this odds ratio is significant within the enriched area, as indicated in **Fig. 32** (lower right panel), suggesting that the reproducibility between replicates is statistically viable, particularly in introns, 3'UTR, and CDS.

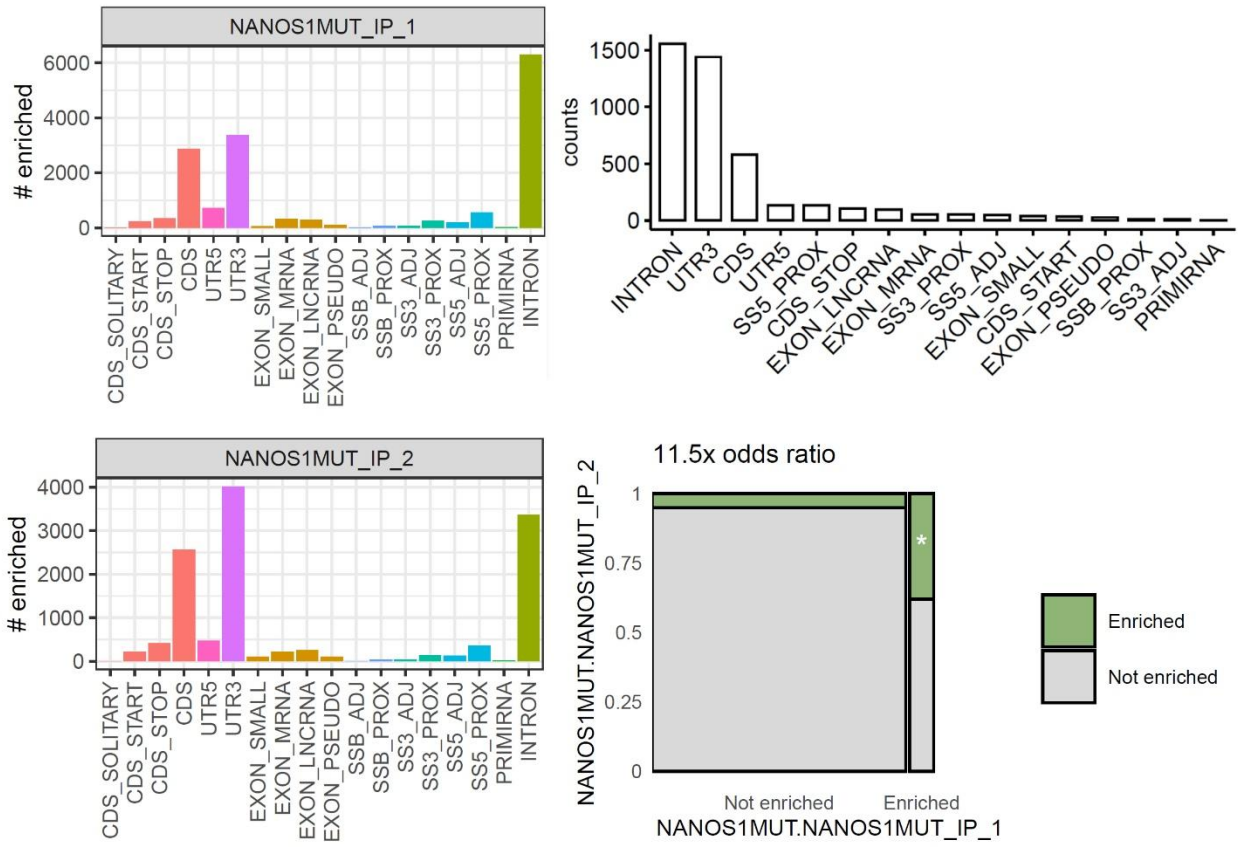


Figure 32. Enrichment of MUT-NANOS1 binding across target RNA sites from two independent eCLIP pre-me replicates. The analysis entails RNA-binding enrichment of immunoprecipitated MUT-NANOS1 from two independent replicates: NANOS1 MUT_IP_1 (top left panel) and NANOS1 MUT_IP_2 (bottom left panel). RNA regions are categorized into feature types such as CDS, UTRs, exons, introns, small RNAs, and SS'. The top right panel displays the counts of enriched bound RNA sites from both replicates, highlighting the top three most enriched RNA features. The odds ratio plot (bottom right panel) compares the enrichment across the two immunoprecipitated replicates of the MUT-NANOS1 protein to assess reproducibility in the binding pattern. In this plot, the green region represents RNA features that are significantly enriched in both replicates (* denotes significance).

4.4 Overlap analysis of enriched RNA interactome of WT- and MUT-NANOS1 with other RNA binding proteins (RBPs) from ENCODE database

To place the most enriched binding sites of WT- and MUT-NANOS1 RNA interactomes in the context of other commonly analysed RBPs via eCLIP, a comprehensive overlap analysis was performed using eCLIP data from previously studied RBPs in HepG2 and K562 cell lines, as provided by the ENCODE database (https://www.encodeproject.org/). The results are depicted in **Fig. 33A**, where a t-distributed stochastic neighbor embedding (t-SNE) plot depicts each RBP as a single point color-coded based on its primary RNA-binding target. Notably, MUT-NANOS1 is positioned at the top of the plot, clustering with RBPs that preferentially bind 3′ UTRs (light blue), whereas WT-NANOS1 clusters near the bottom with RBPs that primarily bind intronic regions (dark green). This distinct clustering corroborates the observation that MUT-

NANOS1 has shifted its binding preference away from intronic regions toward the 3'UTR, while WT-NANOS1 is predominantly binding intronic sites based on its alignment with other well-studied RBPs.

Furthermore, in **Fig. 33B**, the findings from **Fig. 31** and **Fig. 32** are consolidated into stacked bar plots comparing binding-site distributions for WT- and MUT-NANOS1 as fractions of their total binding sites. Notably, WT-NANOS1 binding is predominantly localized to intronic regions, which constitute the largest fraction of total binding sites (**Fig. 33B**), suggesting a potential role in pre-mRNA processing, intron-associated RNA stability or other nuclear functions. In contrast, MUT-NANOS1 displays a substantial shift in binding preferences with reduced intronic binding and a pronounced increase in the 3'UTR and CDS regions **(Fig. 33B)**. This shift indicates altered RNA-binding specificity, favoring mature mRNAs rather than pre-mRNA elements. The elevated 3′ UTR binding suggests a potential gain-of-function in post-transcriptional gene regulation, and the increased CDS binding points to interactions with actively translated transcripts that may affect mRNA stability or protein synthesis. Furthermore, statistical analysis revealed a highly significant difference between the WT- and MUT- binding site distributions (**Fig. 33B**, $p < 2.2 \times 10^{-16}$), confirming that the observed changes are unlikely to be due to random variation.

Lastly, the eCLIP data in **Fig. 33B** are consistent with the overlap analysis shown in **Fig. 33A**. Together, the t-SNE plot and stacked bar plots confirm that, at the pre-me stage, WT-NANOS1 strongly prefers intronic binding, whereas MUT-NANOS1 progressively shifts toward the 3′ UTR and CDS, implying a reallocation of its functional role.

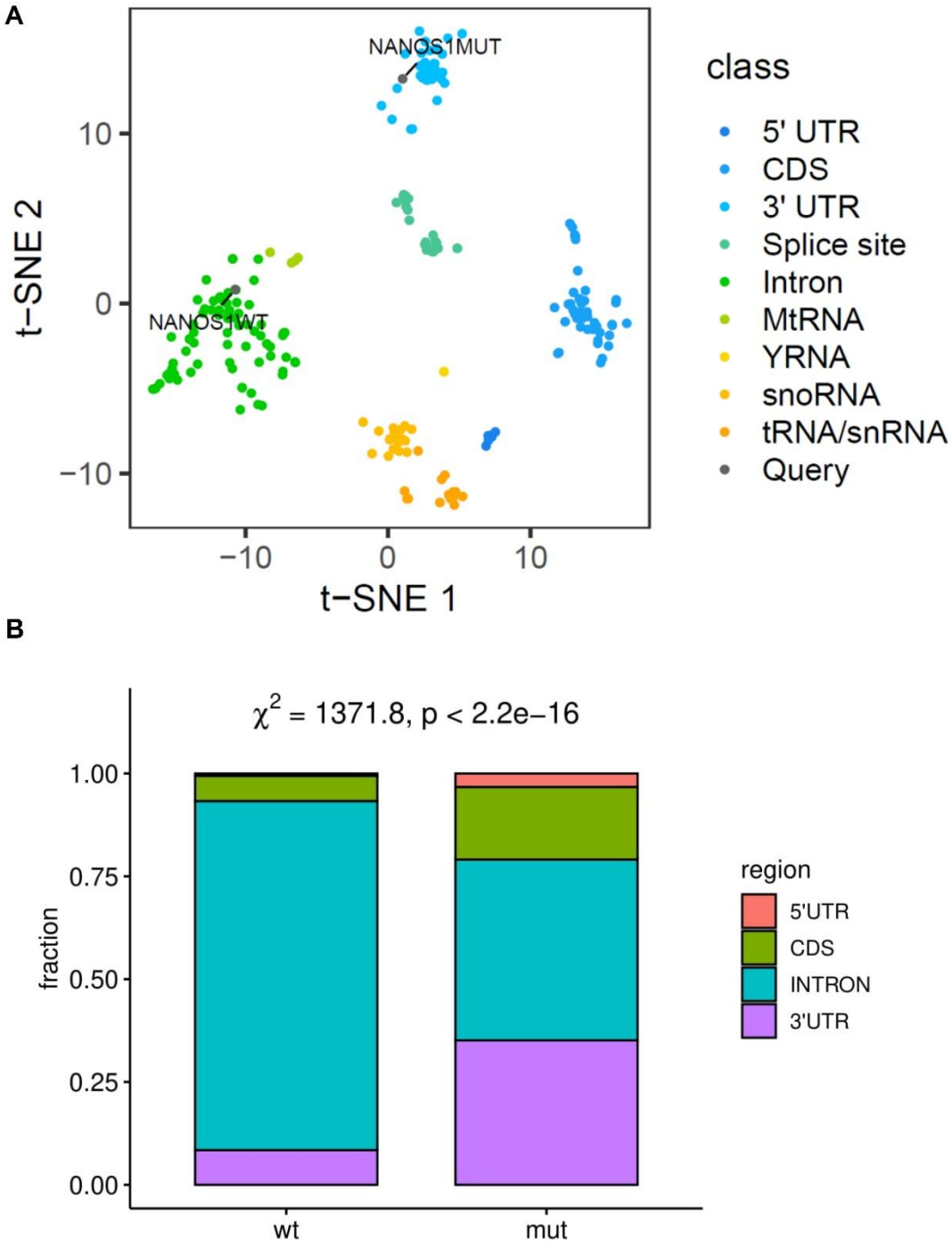


Figure 33. Comparison of RNA binding specificity between MUT- and WT-NANOS1. A- A t-SNE plot illustrating the clustering of RNA-binding classes for MUT-NANOS1 and WT-NANOS1, based on the eCLIP data from **Figs 31** and **32**. Each point represents an RNA-binding protein (RBP) from eCLIP experiments uploaded to the ENCODE database using HepG2 and K562 cell lines. The RBPs are color-coded by RNA feature type, including 5′ UTR, CDS, 3′UTR, splice site, intron, mitochondrial RNA (MtRNA), Y RNA (YRNA), small nucleolar RNA (snoRNA), and transfer RNA/small nuclear RNA (tRNA/snRNA). **B**- The stacked bar plot compares the distribution of NANOS1 binding sites in WT- and MUT-NANOS1 RNA interactomes (x-axis) across different genomic regions: 5'UTR, CDS, INTRON, and 3'UTR. The fractions (y-axis) represent the proportion of binding sites in each region relative to the total number. A chi-square test ($\chi^2 = 1371.8$, $p < 2.2e^{-16}$) indicates a highly significant difference between the distributions of binding sites in WT- and MUT-NANOS1 RNA targets.

4.5 The MUT-NANOS1 interactome encompasses a significantly higher number of mRNAs than WT-NANOS1

An overlap analysis was conducted to merge the binding sites and bound genes potentially by either WT-NANOS1 and/or its MUT counterpart, as illustrated in the Venn diagrams in **Fig. 34A-B**. The analysis revealed that WT-NANOS1 possesses 2,341 unique binding sites, whereas MUT-NANOS1 has 2,942 unique sites, with 1,388 binding sites shared between the two counterparts (**Fig. 34A**). Despite the comparable number of binding sites, the MUT-NANOS1 binds to a significantly higher number of genes: 810 unique genes compared to only 164 unique genes bound by WT-NANOS1, with 364 genes shared between them (**Fig. 34B**). This increase in gene targets, despite similar binding site overlap, suggests that MUT-NANOS1 exhibits a broader and less selective RNA-binding profile, potentially indicating a loss of specificity. The MUT protein's shift toward binding a larger variety of genes may reflect an alteration in its RNA-regulatory function, contributing to broader and possibly dysregulated RNA interactions. Furthermore, the overlap analysis (**Fig. 34A–B**) corroborates the previously observed differences in RNA-binding preferences between WT- and MUT-NANOS1 (**Figs. 31**, **32**, and **33A–B**). Unlike WT-NANOS1, MUT-NANOS1 displays a broader binding preference, not only associating with introns but also showing increased enrichment in 3' UTR and CDS. This shift in binding targets likely explains the substantial increase in the number of genes (810 unique genes, **Fig. 34B**) enriched. The MUT-NANOS1 protein appears to have lost intronic-binding specificity and adopted a broader, less selective profile across a wider array of mRNAs, expanding its regulatory scope at the potential cost of reduced specificity and leading to dysregulated RNA processing and gene expression in pre-me cells.

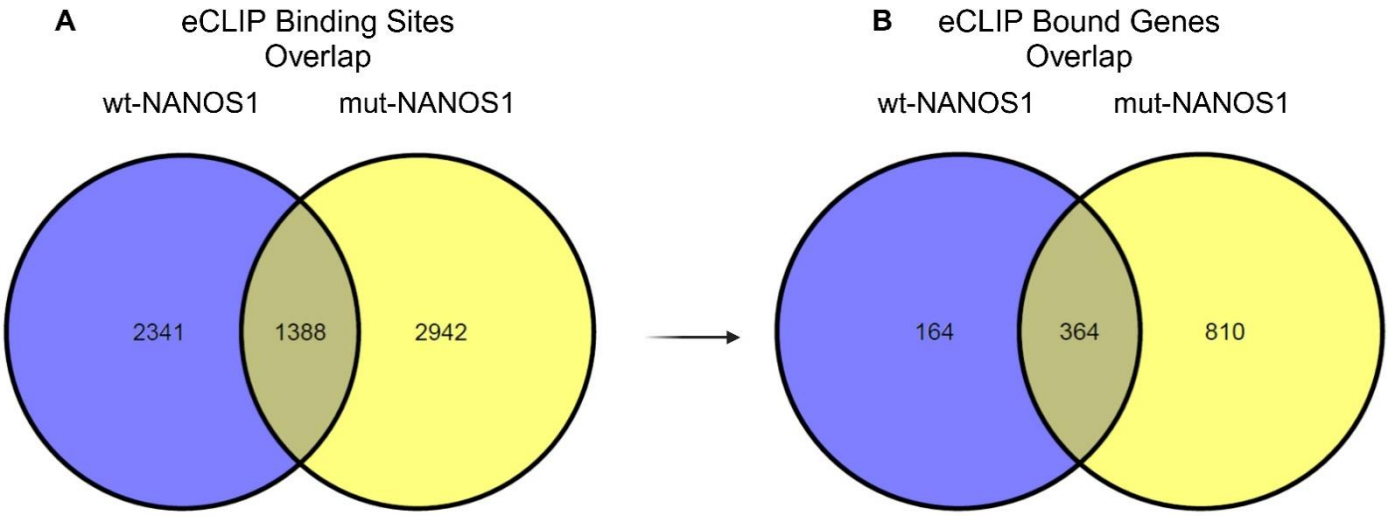


Figure 34. Comparative eCLIP analysis of WT- and MUT-NANOS1 reveals a larger gene-encoded RNA interactome for MUT-NANOS1. The two Venn diagram sets depict comparison of the eCLIP binding profiles of WT-NANOS1 and MUT-NANOS1. **A**-Venn Diagram: Illustrates the overlap in eCLIP binding sites between WT-NANOS1 and MUT-NANOS1. **B**- Venn Diagram: Depicts the overlap in genes bound by WT-NANOS1 and MUT-NANOS1, as determined from the locations of their binding sites.

5. Dynamic RNA interactome alterations in WT- and MUT-NANOS1 overexpression based on RNA-seq. and eCLIP overlap

The eCLIP assay was applied in conjunction with bulk RNA sequencing to investigate the differential gene expression (DGE) of NANOS1-bound mRNAs during different stages of stem cell differentiation. Additionally, overlap of data from both assays aimed to highlight the dynamic alterations in the potential RNA interactome bound by the respective NANOS1-overexpressed proteins. Unlike eCLIP, bulk RNA sequencing was performed on three cell lines: unedited control W15, WT-NANOS1 1.2.5.1, and MUT-NANOS1 3.2.5.2, at various differentiation stages, including pre-me and Day 4 PGC sorted cells. Detailed methods of sample collection and sorting are provided in the Methods section (**Fig. S10–S13**).

5.1 Minimal effect of dox on control cell line W15 at pre-me and PGC stages

Given that dox was used to induce NANOS1 overexpression in the cell lines, its effect alone was initially assessed in the unedited control cell line W15 (**Fig. 35A–D**). Three biological replicates were compared before and after dox addition at both the pre-me stage and the Day 4 sorted PGC population (PG). Principal component analysis (PCA) plots (**Fig. 35A** and **35C**) illustrate the correlation between replicates of W15 cells at the aforementioned differentiation stages before and after dox treatment. In the pre-me stage, the first principal component (PC1) accounts for 34% of the variance between samples due to dox addition, while the second principal component (PC2) accounts for 27% of the variance within replicates (**Fig. 35A**). Similarly, in the PGC stage, PC1 represents 30% of the variance attributable to dox treatment, and PC2 accounts for 21% of the variance within replicates (**Fig. 35C**).

The clustering of samples along PC1 and PC2 in both differentiation stages indicates that the overall transcriptional profiles are relatively similar, suggesting that dox induces only minor changes in gene expression in pre-me and PGC sorted cells. This observation is further supported by the mean average (MA) plots representing the fold changes of genes before and after dox addition in both pre-me (**Fig. 35B**) and PGCs (**Fig. 35D**). The majority of genes align along the center line (log fold change close to zero), indicating minimal changes in gene expression following dox treatment. While a few genes exhibit slight upregulation or downregulation, these changes are not extensive (**Fig. 35B** and **35D**). Therefore, both the PCA and MA plots suggest that dox treatment induces only slight alterations in the gene expression profiles of W15 pre-me and PGCs.

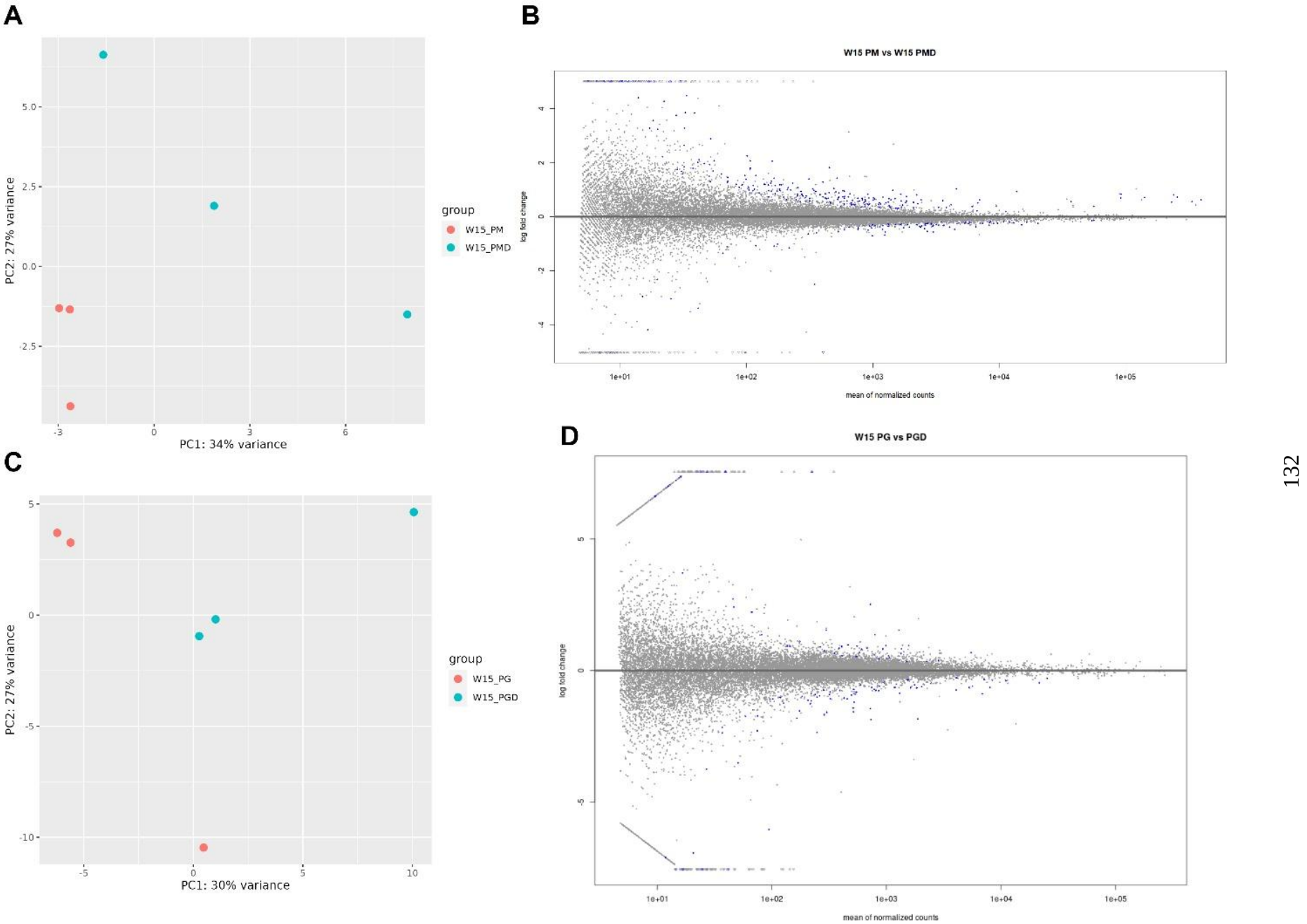

A
PC2: 27% variance
PC1: 34% variance
group
W15_PM
W15_PMD
B
W15 PM vs W15 PMD
log fold change
mean of normalized counts
C
PC2: 27% variance
PC1: 30% variance
group
W15_PG
W15_PGD
D
W15 PG vs PGD
log fold change
mean of normalized counts

Figure 35. Comparison of control cell line W15 before and after dox addition at the pre-me and PGC differentiation stages. A- Principal component analysis (PCA) of W15 pre-me biological replicates before (W15-PM) and after (W15-PMD) dox addition at 500 ng/ml. PC1 represents the component with the highest variance (between samples), and PC2 represents the component with the second highest variance (between replicates of the same sample). **B**- MA plot following differential gene expression analysis between W15-PM and W15-PMD. The x-axis represents the mean of normalized counts, indicating the average expression level of genes across the two conditions. The y-axis represents the log2 fold change per gene between the two conditions. Each point represents a gene: grey points denote genes that are not significantly differentially expressed, and blue points denote genes with significant differential expression based on the adjusted p-value. **C**- PCA of W15 PGC biological replicates before (W15-PG) and after (W15-PGD) dox addition. **D**- MA plot visualizing differential gene expression analysis between W15-PG and W15-PGD replicates.

5.2 Gene expression variance between control W15 and NANOS1 cell lines WT 1.2.5.1 and MUT 3.2.5.2 prior to dox addition

Prior to comparing post-dox transcriptional profiles between the control cell line W15 and the NANOS1 cell lines, PCA plots (**Fig. 36A-D**) were generated to assess baseline transcriptional variance. These plots illustrate the variance in gene expression between the control cell line and edited ones before the overexpression of either MUT- or WT-NANOS1, where ideally the generated clones should exhibit similar transcriptional profiles. At the pre-me stage (**Fig. 36A-B**), the MUT-NANOS1 cell line showed 52% variance along PC1 relative to W15, with 21.6% variance within the replicates (PC2, **Fig. 36A**), indicating slightly higher separation than expected, likely due to genetic modifications introduced during PiggyBac editing. In contrast, the WT-NANOS1 cell line exhibited a 75% variance along PC1 compared to W15, with only 9% within-replicate variance (PC2, **Fig. 36B**). This was an unexpected finding, as WT-NANOS1 and W15 pre-me cells were anticipated to be similar at baseline; however, it aligns with immunofluorescence findings, where WT 1.2.5.1 cells exhibited lower pre-me marker (EOMES and Brachyury T) expression before dox induction compared to both MUT 3.2.5.2 and W15 cells (**Fig. 20A–C** and **Fig. 23A–C**).

At the PGC stage (**Fig. 36C–D**), the MUT-NANOS1 cell line again diverged from W15 by 60% along PC1, with 16% variance within replicates (**Fig. 36C**), mirroring the slightly higher-than-expected difference attributable to genetic editing or extended culturing and subsequent cell sorting into PGCLCs. In contrast, the WT-NANOS1 cell line differed by 93% along PC1, with only 4% within-replicate variance (**Fig. 36D**); removal of an outlier replicate left only two WT replicates and two W15 replicates. This extreme divergence suggests fundamental differences in transcriptional profiles that preclude conclusive comparisons between the WT-NANOS1 cell line and W15 or within the WT-NANOS1 line pre- and post-dox induction. These findings suggest that the WT pre-me and PGC cells, prior to dox induction, may have inherent baseline transcriptional differences due to potentially lower baseline pluripotency. This is further supported by the lower immunofluorescence levels of pluripotency markers (OCT4, NANOG, and SOX2) in WT 1.2.5.1 stem cells compared to W15 control (**Fig. 18**), likely reflecting inherent genetic or regulatory changes acquired during cell line generation. It is highly probable that these factors contribute to the divergence of the cell lines even before NANOS1 overexpression, with the WT-NANOS1 cell line showing unexpectedly large deviations (**Fig. 36B-D**) from the W15 control across differentiation.

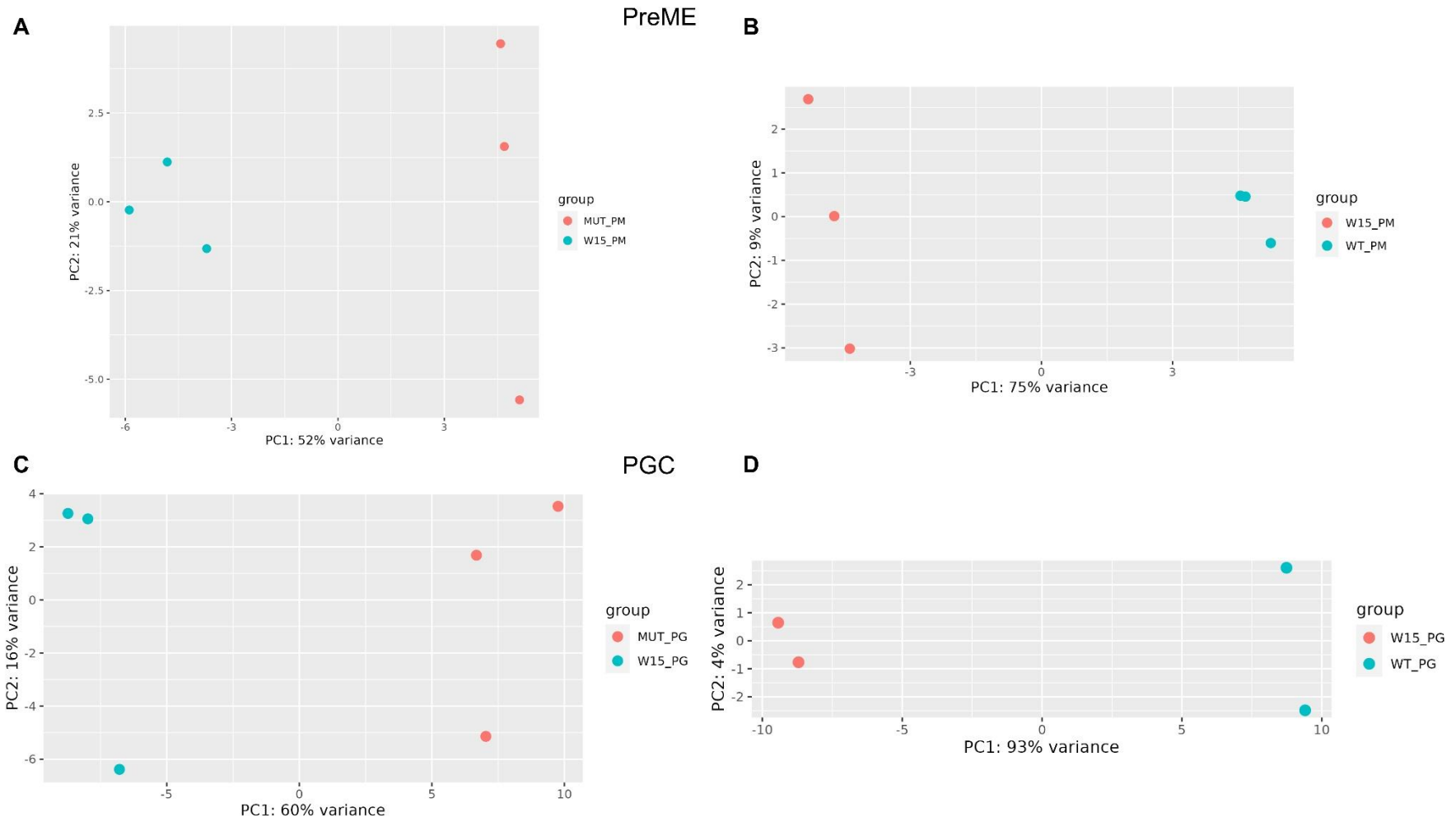

PreME
A
PC2: 21% variance
PC1: 52% variance
group
MUT_PM
W15_PM
B
PC2: 9% variance
PC1: 75% variance
group
W15_PM
WT_PM
PGC
C
PC2: 16% variance
PC1: 60% variance
group
MUT_PG
W15_PG
D
PC2: 4% variance
PC1: 93% variance
group
W15_PG
WT_PG

Figure 36. Principal component analysis (PCA) of pre-me and PGC differentiation stage comparing control cell line W15 with WT- and MUT-NANOS1 cell lines prior to dox induction (baseline state). A- PCA plot comparing pre-me replicates of W15 and the WT cell lines. **B**- PCA plot comparing pre-me replicates of W15 and the MUT cell lines. **C**- PCA plot comparing PGC replicates of W15 and the MUT cell lines. **D**- PCA plot comparing PGC replicates of W15 and the WT cell lines. In each PCA plot, the first principal component (PC1) represents the axis of greatest variance between the cell lines, while the second principal component (PC2) represents the second greatest variance, typically reflecting variability within replicates.

5.3 Gene expression variance in WT-NANOS1 1.2.5.1 cell line following dox induction at the pre-me and PGC differentiation stages

Due to the observed differences between the cell lines prior to dox treatment (**Fig. 36A-D**), comparisons were focused within each cell line before and after dox induction to assess the effect of NANOS1 overexpression. In the WT cell line at the pre-me stage (**Fig. 37A** and **Fig. 37C**), PCA revealed a 46% variance along PC1 was observed between dox induced (WT-PMD) and uninduced (WT-PM) samples, driven largely by one induced replicate, suggesting modest transcriptional shifts. At the PGC stage, (**Fig. 37B** and **Fig. 37D**) the variance increased to 60% along PC1 between PGCs before (WT-PG) and after dox induction (WT-PGD), with a 25% variance along PC2 within replicates of the same condition. The more pronounced separation following NANOS1 overexpression indicates that transcriptional changes are more significant at this stage compared to pre-me. This suggests that NANOS1 overexpression in PGCs may further fine-tune gene expression, including the core transcriptional program established during the earlier pre-me stage.

Gene set enrichment analysis (GSEA) visualized via UpSet plots (**Fig. 37C**) at the pre-me stage showed upregulation in spliceosome-related processes, including spliceosomal small nuclear ribonucleoprotein (snRNP) complex and its assembly, mRNA cis splicing via the spliceosome, and pre-mRNA binding (**Fig. 37C**). This enhancement of splicing-related pathways and pre-mRNA binding is consistent with eCLIP findings demonstrating NANOS1's binding to intronic regions (**Fig. 31**). The substantial positive fold changes suggest that overexpressed NANOS1 may modulate spliceosome assembly and splicing machinery, which are critical for RNA processing as cells transition toward germ cell fate. Conversely, there is a concurrent downregulation of metabolic pathways, indicating decreased enrichment in processes such as oxygen metabolic activity and nucleotidase activity (**Fig. 37C**). This reduction may be associated with cellular differentiation toward PGC specification, where NANOS1 may downregulate other metabolic processes in favor of RNA processing.

By the PGC stage (**Fig. 37D**), there was an overall trend of suppressed somatic developmental pathways, including heart development, sensory organ development, muscle tissue, and nephron epithelium development. This downregulation of somatic developmental genes aligns with NANOS1's role in PGC specification, where it is expected to suppress somatic pathways to maintain germ cell identity. The modest enhancement of spliceosomal components (U2 snRNP, **Fig. 37D**) suggests that NANOS1 may continue to be involved in splicing-related functions during the PGC stage, albeit to a lesser extent. Additionally, downregulation of lipid digestion and cholesterol import processes imply that these metabolic activities are

less central to PGC differentiation and more characteristic of somatic lineages. Notably, the acrosomal membrane category exhibited a positive fold change (**Fig. 37D**) suggesting a function in later sperm development and maturation, highlighting its importance in later, sex-specific stages of germ cell differentiation.

Lastly, the receptor for advanced glycation end products (RAGE) binding category showed a marked rise, albeit with high variability (**Fig. 37D**). RAGE is a multiligand receptor involved in various physiological and pathological processes, including inflammation, cellular stress responses, and immune regulation. The receptor binds advanced glycation end products (AGEs) and other ligands such as S100 proteins (Hudson and Lippman, 2017). The increase in RAGE receptor binding may indicate a stress response in PGCs due to NANOS1 overexpression, as the artificial overexpression could induce cellular stress. Overall, these data indicate that NANOS1 overexpression promotes germline transcriptional programs and represses somatic pathways, likely by fine-tuning RNA processing critical for early germ cell fate.

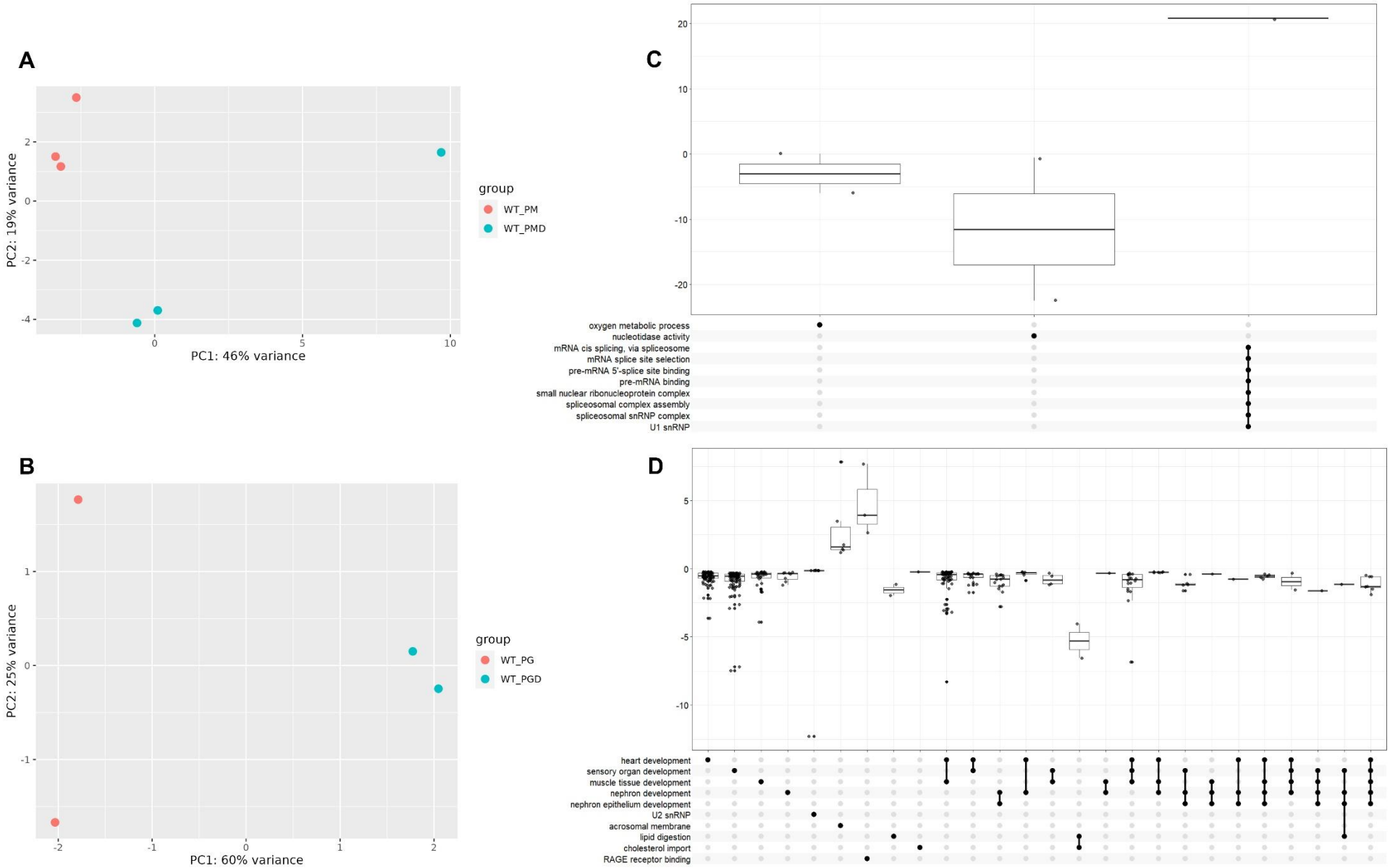

A
PC2: 19% variance
PC1: 46% variance
group
WT_PM
WT_PMD
B
PC2: 25% variance
PC1: 60% variance
group
WT_PG
WT_PGD
C
oxygen metabolic process
nucleotidase activity
mRNA cis splicing, via spliceosome
mRNA splice site selection
pre-mRNA 5'-splice site binding
pre-mRNA binding
small nuclear ribonucleoprotein complex
spliceosomal complex assembly
spliceosomal snRNP complex
U1 snRNP
D
heart development
sensory organ development
muscle tissue development
nephron development
nephron epithelium development
U2 snRNP
acrosomal membrane
lipid digestion
cholesterol import
RAGE receptor binding

Figure 37. Principal component analysis (PCA) and visualization of gene set enrichment analysis (GSEA) via UpSet plots in WT 1.2.5.1 cell line across developmental stages pre-me and PGC. A- PCA plot of pre-me WT 1.2.5.1 replicates before (WT-PM) and after dox addition (WT-PMD), which induces overexpression of WT-NANOS1. The first principal component (PC1) represents the highest variance between samples, and the second principal component (PC2) represents the second highest variance within conditions. **B**- UpSet plot visualizing GSEA results of the differentially expressed genes between WT-PM and WT-PMD conditions. The bottom portion of the UpSet plot represents the intersections between different GSEA pathways (biological processes). Each column corresponds to a specific intersection or overlap between GSEA terms, indicated by black dots in the matrix. A single dot represents unique terms, while multiple connected dots represent intersections among multiple terms. The upper portion features boxplots depicting the distribution of fold changes of the intersecting genes, including statistics such as range, median, and variability for each intersection of GSEA terms. **C**- PCA plot of PGC WT replicates before (WT-PG) and after dox addition (WT-PGD), inducing overexpression of WT-NANOS1 in PGC cells. **D**- UpSet plot visualizing GSEA results of the differentially expressed genes between WT-PG and WT-PGD conditions, following the same structure as in panel B.

5.4 Gene expression variance and altered pathways in MUT-NANOS1 3.2.5.2 cell line following dox induction at the pre-me stage

In addition to examining WT-NANOS1, the MUT-NANOS1 3.2.5.2 cell line was analysed before (MUT-PM) and after (MUT-PMD) dox induction at the pre-me stage (**Fig. 38A-B**). PCA (**Fig. 38A**) revealed a 41% variance along PC1 between MUT-PM and MUT-PMD, with 25% variance along PC2 within replicates, indicating distinct transcriptional changes induced by MUT-NANOS1 overexpression.

The heatmap derived from DGE between the two conditions (**Fig. 38B**) further demonstrates an appropriate grouping of replicates based on their transcriptional profiles. Notably, an extensive downregulation (blue clusters) is depicted, suggesting broad repression of key genes linked to primed pluripotency, including CD24 (Collier et al., 2017) and HPAT5, a key component of the pluripotency network, interacting with the let-7 microRNA family (Durruthy-Durruthy et al., 2016). This repression suggests that MUT-NANOS1 overexpression may compromise essential processes for pluripotency of pre-me cells. Furthermore, ZIC2, which is upregulated during PGC specification trajectory starting with pre-me cells as part of the MAPK pathway and later downregulated by PRDM1 in PGCs (Tang et al., 2022), is among the genes downregulated following MUT-NANOS1 overexpression (**Fig. 38B**). This indicates not only a reduction in cell propagation but also a potential decrease in PGC competence.

Collectively, these findings indicate that MUT-NANOS1 overexpression broadly suppresses essential pluripotency and early germ cell specification pathways in pre-me cells, potentially undermining their capacity to progress toward PGC fate.

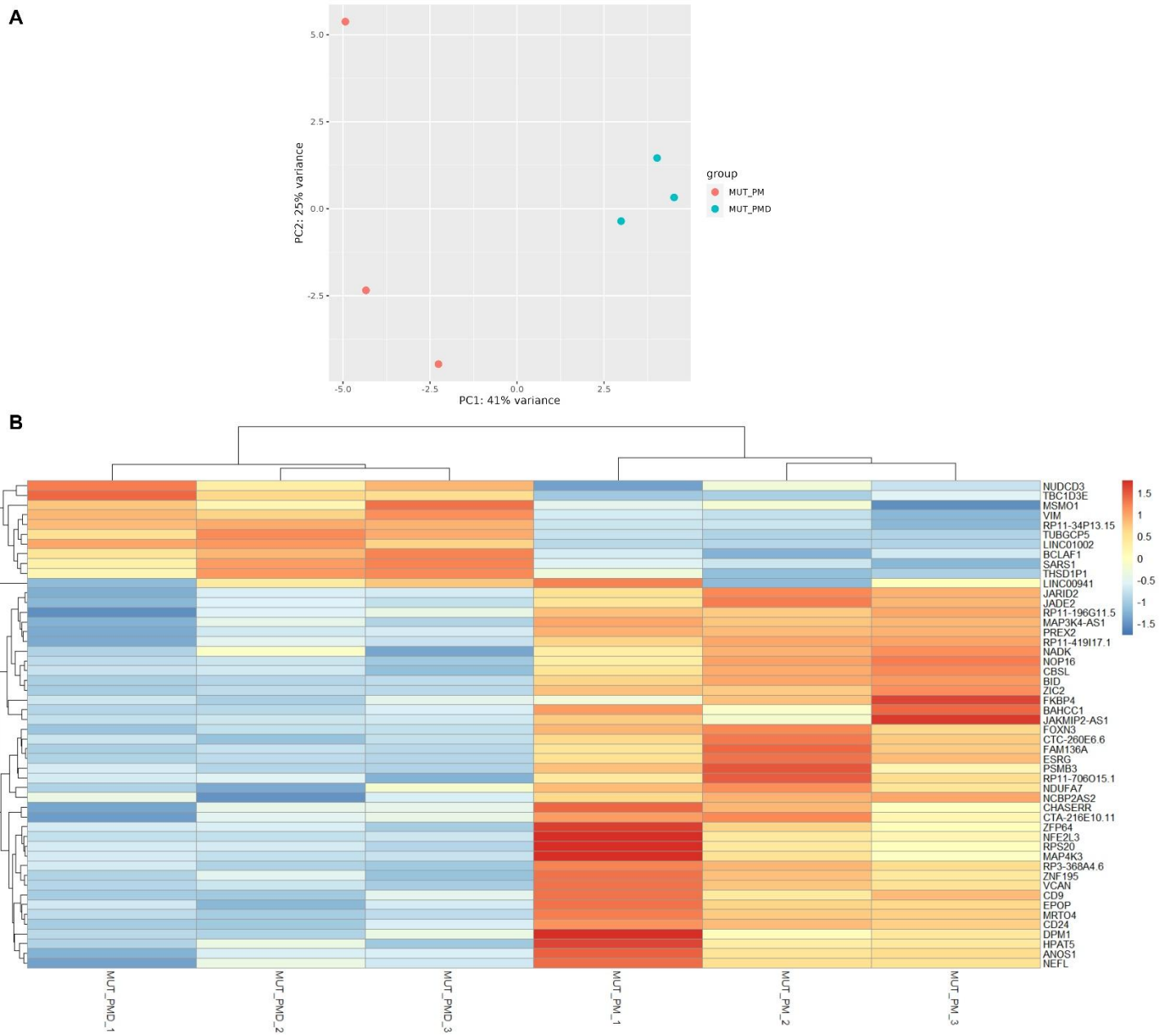


Figure 38. PCA plot and heatmap of gene expression in MUT-NANOS1 3.2.5.2 cell line at the pre-me differentiation stage. A- Principal Component Analysis (PCA) plot illustrating the separation between MUT cell line pre-me replicates before (MUT-PM, red) and after dox addition (MUT-PMD, teal), which induces overexpression of MUT-NANOS1 in pre-me cells. **B**- Heatmap depicting the expression levels of differentially expressed genes across three biological replicates for each condition: MUT-PM and MUT-PMD. Red indicates upregulated genes, while blue denotes downregulated genes. The heatmap is hierarchical, with both genes and samples clustered based on their expression patterns. Samples are organized along the bottom axis, and genes are clustered along the left side. Similar expression patterns among samples are grouped closer together, as indicated by the dendrograms at the top (for samples) and left side (for genes) of the heatmap.

5.4.1 The top 10 downregulated pathways following MUT-NANOS1 overexpression in pre-me include WNT signaling, stem cell differentiation, and protein synthesis in the cytoplasm

Based on the differential expression analysis, which revealed a predominance of repressed genes, as visualized in the heatmap (**Fig. 32B**), gene ontology (GO) analysis was performed on the downregulated genes to elucidate the affected pathways following overexpression of MUT- NANOS1 pre-me cells. The top ten significantly downregulated pathways (**Fig. 39**) include essential biological processes related to translation, stem cell maintenance and differentiation, and signaling pathways such as WNT (wingless/INT-1), which is crucial for cell-cell communication and development. The two most significantly downregulated processes are "cytosolic ribosome" and "cytoplasmic translation," both directly correlated with protein synthesis and indicating diminished translational machinery.

Regarding the downregulation of "stem cell differentiation" and "stem cell population maintenance," it is important to note that the pre-me stage represents a transient phase clustering with hESCs and mesendoderm (ME), whereas human PGCLCs and human PGCs cluster separately (Tang et al., 2022). Therefore, transcriptomically, pre-me cells lie between stemness and differentiation, both processes that are significantly negatively affected post MUT-NANOS1 overexpression (**Fig. 39**). Maintaining a balance between stem cell maintenance and progression toward early ME is crucial during the pre-me stage (Tang et al., 2022), and any deviation can impact the cell's priming toward the desired lineage.

5.4.1.1 WNT signaling pathway

Another significantly affected pathway is the WNT signaling pathway and its associated impacts on cell-cell signaling, specifically via its canonical form (**Fig. 39**). *In vitro* differentiation of hESCs into ME relies on canonical WNT and Activin/Nodal signaling, with an epigenetic enhancer cluster (including EOMES) activating WNT-mediated processes critical for both ME formation and PGC competence (Tang et al., 2022). Furthermore, the study by Tang et al. (2022) highlighted the role of the canonical WNT signaling pathway, particularly mediators like LEF1, TCF3, and TCF7L2, in activating early ME enhancers during human PGC specification. Therefore, the downregulation of WNT signaling significantly affects both the priming process at the pre-me stage and the acquisition of competence for PGC specification, an effect influenced by MUT-NANOS1 overexpression.

5.4.1.2 Pathways related to mesenchyme development and mesenchymal cell differentiation

Processes such as "mesenchyme development" and "mesenchymal cell differentiation" are also highlighted, reflecting an affected developmental trajectory of cells during the pre-me transition (**Fig. 39**). Studies in mice have shown that nascent mesoderm fails to delaminate and migrate away from the primitive streak due to the inability to downregulate E-cadherin and undergo epithelial-to-mesenchymal transition (EMT) (Arnold et al., 2008). Experimental work by Arnold et al. (2008) demonstrated that EOMES, a downstream target of canonical WNT, plays pivotal roles in anterior-posterior axis patterning, EMT, and definitive endoderm

specification. Incorrect patterning and migration of germ layers can significantly affect not only proper germ layer formation but also the correct specification of PGCs *in vivo*.

Overall, the fact that pathways related to ME development, WNT signaling, and stem cell differentiation are all downregulated suggests a broader disruption in the normal timing of differentiation events. While WT-NANOS1 appears to be more involved in pre-mRNA regulation during pre-me, the MUT counterpart loses this ability and shifts toward binding and regulation of mature mRNAs, thereby misregulating the timing of cell specification processes. Consequently, MUT-NANOS1 interferes with normal differentiation cues at the pre-me stage, delaying or altering both ME and PGC trajectories, highlighting the MUT-NANOS1 protein's misregulation of essential developmental processes.

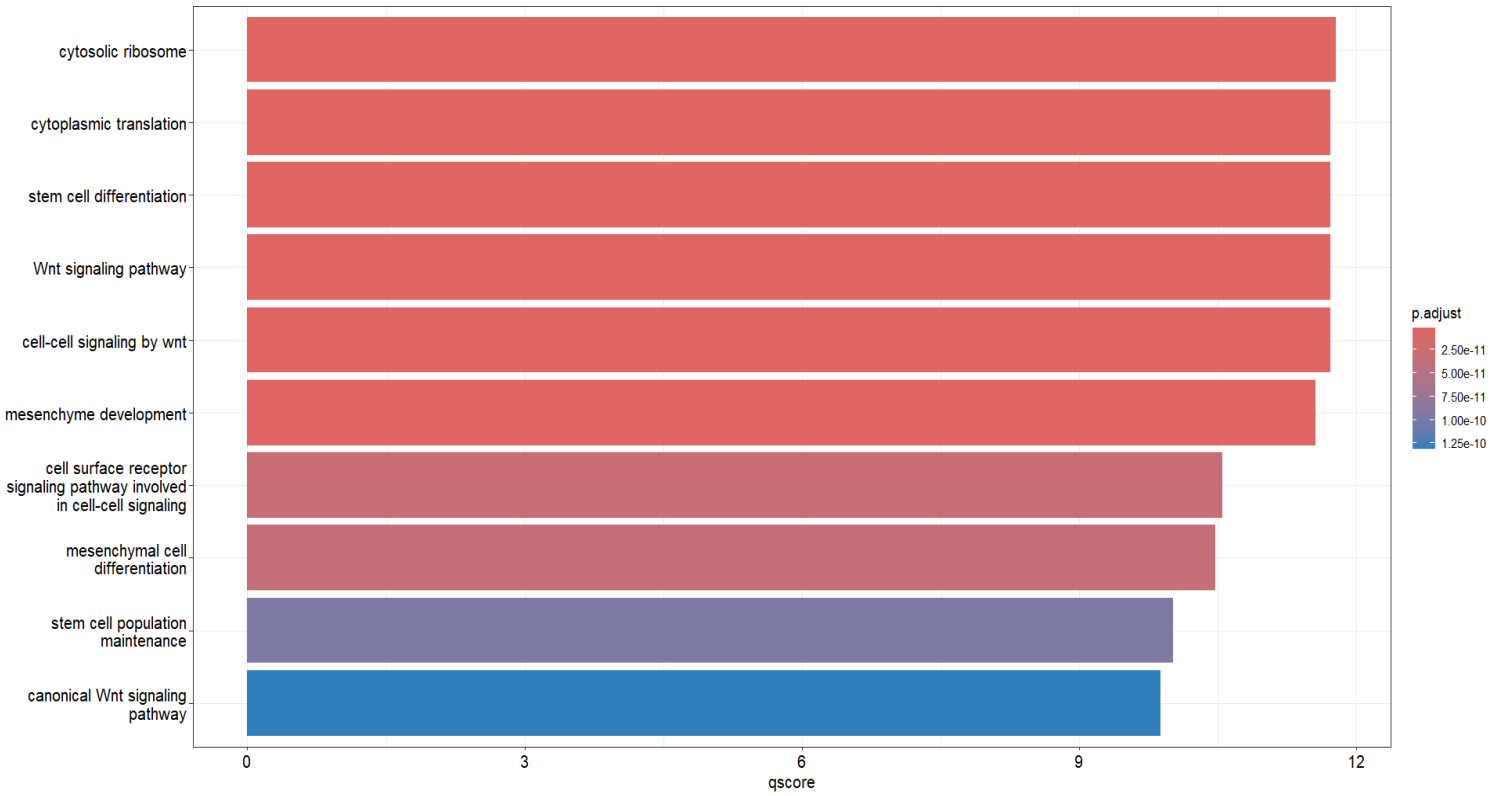


Figure 39. Top 10 downregulated pathways from gene ontology (GO) enrichment analysis following differential expression in MUT expressing pre-me cells. This figure illustrates the top ten downregulated pathways identified from GO enrichment analysis after performing differential expression analysis between MUT pre-me cells before (MUT-PM) and after overexpression of MUT-NANOS1 in MUT 3.2.5.2 cells (MUT-PMD). The differential expression analysis was conducted following the removal (blacklisting) of genes affected solely by dox induction. X-axis (q-score): Represents the significance level (adjusted p-value) of the downregulation, with higher q-scores indicating stronger statistical significance. Y-axis: Lists the names of the downregulated pathways, representing key biological processes affected by MUT-NANOS1 overexpression. Colour Gradient: Depicts the adjusted p-value (p.adjust) for each pathway. Darker blue indicates lower statistical significance, while red signifies higher statistical significance.

5.5 Gene expression variance and altered pathways in MUT-NANOS1 3.2.5.2 cell line following dox induction at the PGC stage

Building upon the observations at the pre-me stage, a differential expression analysis was conducted to assess the effects of MUT-NANOS1 overexpression at the PGC stage (**Fig. 40A-B**). The resulting PCA plot (**Fig. 40A**) depicts an even higher variance between PGCs post MUT-NANOS1 overexpression (MUT-PGD), and their baseline counterpart replicates (MUT-PG), with PC1 accounting for 74% of the variance. The variance within replicates is lower compared to the pre-me stage, at 11% along the PC2 (**Fig. 40A**). This pronounced separation between MUT-PG and MUT-PGD replicates indicates that dox-induced overexpression of MUT-NANOS1 exerts a significant impact on the transcriptional landscape of PGCs. This differential gene expression profile is further highlighted by the functional enrichment of gene sets, visualized in the UpSet plot following GSEA (**Fig. 40B**). Overrepresented categories with positive fold changes include skeletal system development, extracellular matrix organization, and collagen-containing extracellular matrix (**Fig. 40B**). These findings suggest that genes associated with somatic development, such as skeletal and structural pathways, are being upregulated. In contrast to WT-NANOS1 (**Fig. 37D**), which downregulated non-germline pathways, MUT-NANOS1 appears to lose specificity and activates non-germline gene expression. Additional somatic differentiation pathways that are significantly upregulated involve striated muscle tissue development and muscle organ morphogenesis (**Fig. 40B**), further emphasizing the potential loss of binding specificity of MUT-NANOS1 leading to off-target activation of somatic developmental programs. Moreover, metabolic categories associated with mitochondrial function and energy production, such as cytochrome complexes, cytochrome c oxidase activity, and respiratory chain complexes, are also upregulated (**Fig. 40B**).

This indicates that MUT-NANOS1 may promote increased metabolic activity or mitochondrial function, potentially to meet the energy demands associated with differentiation into non-germline cells. Notably, the only category exhibiting a significant negative fold change is associated with pre-mRNA 5'-splice site binding (**Fig. 40B**). This observation contrasts with findings for WT-NANOS1, where pre-mRNA splicing and overall RNA splicing were positively enriched, particularly at the pre-me stage but also at the PGC stage (**Fig. 37C–D**).

Additionally, while WT-NANOS1 predominantly binds intronic regions, MUT-NANOS1 displays a more balanced distribution across intronic and 3′ UTR sites (**Fig. 33B**), indicating reduced specificity in pre-mRNA processing. This broader binding pattern (**Fig. 34B**) might correlate with the upregulation of somatic developmental processes and associated metabolic changes (**Fig. 40B**), potentially at the expense of germline maintenance during the PGC stage.

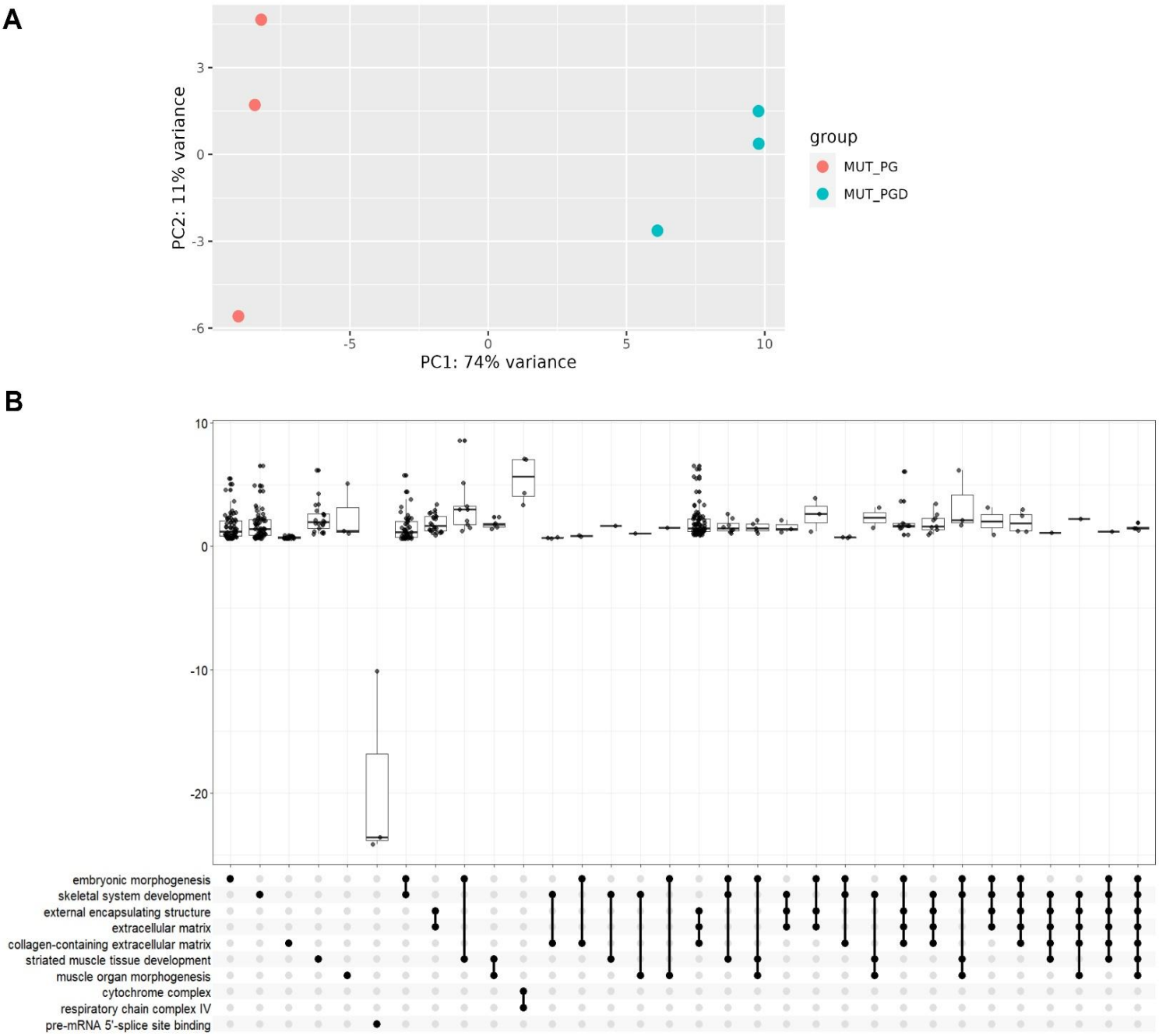


Figure 40. PCA plot and GSEA results visualized *via* UpSet plot showing gene expression and functional enrichment in MUT 3.2.5.2 PGCs. A- The PCA plot visualizes the variance in gene expression between MUT-NANOS1 PGCs before dox induction (MUT_PG, red) and after dox-induced MUT-NANOS1 overexpression (MUT-PGD, teal). The first principal component (PC1) represents the highest variance between the conditions, indicating the impact of MUT-NANOS1 overexpression on the transcriptional landscape. **B**- The UpSet plot displays the intersections of enriched biological processes from the gene set enrichment analysis (GSEA) between the two conditions (MUT-PG vs. MUT-PGD). The intersection matrix (dots at the bottom) indicates which enriched pathways are shared or unique across different sets. Each column represents a specific intersection or overlap between GSEA terms (biological processes), with black dots indicating the involved terms. Single dots represent unique pathways, while connected dots represent shared pathways among multiple sets. The boxplots above the matrix represent the distribution of enriched genes across these sets, providing statistics such as range, median, and variability for each intersection of GSEA terms.

5.5.1 The top 10 upregulated pathways following MUT-NANOS1 overexpression in PGC include upregulated extra cellular matrix, enhanced WNT signaling and embryonic morphogenesis

In line with the GSEA finding visualized *via* the UpSet plot (**Fig. 40B**), GO analysis on significantly upregulated genes (**Fig. 40B** and **Fig. 41**) revealed that MUT-NANOS1 overexpression at the PGC stage activates pathways associated with cell-to-extracellular matrix (ECM) interactions, WNT signaling, developmental processes via embryonic morphogenesis, stress response and cellular remodeling, as indicated by upregulation of lysosomal lumen and endoplasmic reticulum lumen pathways (**Fig. 41**). In the context of normal PGC developmental trajectories, PRDM1 suppresses these processes to facilitate PGC development (Tang et al., 2022), repressing genes like EOMES and ZIC2/3/5 involved in WNT signaling and embryogenesis, as well as the pluripotency regulator SOX2.

Based on the GO findings, MUT-NANOS1 overexpression appears to disrupt the expected PGC developmental process by impeding the proper repression of pathways controlled by PRDM1, such as WNT signaling (**Fig. 41**). By failing to maintain PRDM1-mediated repression, MUT-NANOS1 drives cells to upregulate adhesion, morphogenesis, and WNT pathway, potentially diverting differentiating cells into mesoderm or neurogenesis-like lineages and destabilizing their PGC identity. This could explain the increased ECM interactions and focal adhesion, indicating that the cells are inappropriately and excessively engaging with their surrounding environment within the EB. Such behavior may affect the PGCs' proper migration and integration into correct locations or may signify that the cellular niche more closely resembles a mesoderm-like layer rather than the PGC niche. Similarly, the increased embryonic morphogenesis (**Fig. 41**), which is typically repressed by PRDM1 (Tang et al., 2022), suggests that MUT-NANOS1 may be causing cells to activate non-germ cell fates, such as differentiation into somatic lineages. This could result in a population of cells that fail to remain on the path to becoming PGCs, instead diverting to other lineages.

Lastly, the accompanying increase in endoplasmic reticulum and lysosomal lumen markers may reflect stress responses or protein misfolding, further compromising PGC development. Overall, MUT-NANOS1 disrupts essential PGC specification by misregulating key developmental and stress pathways.

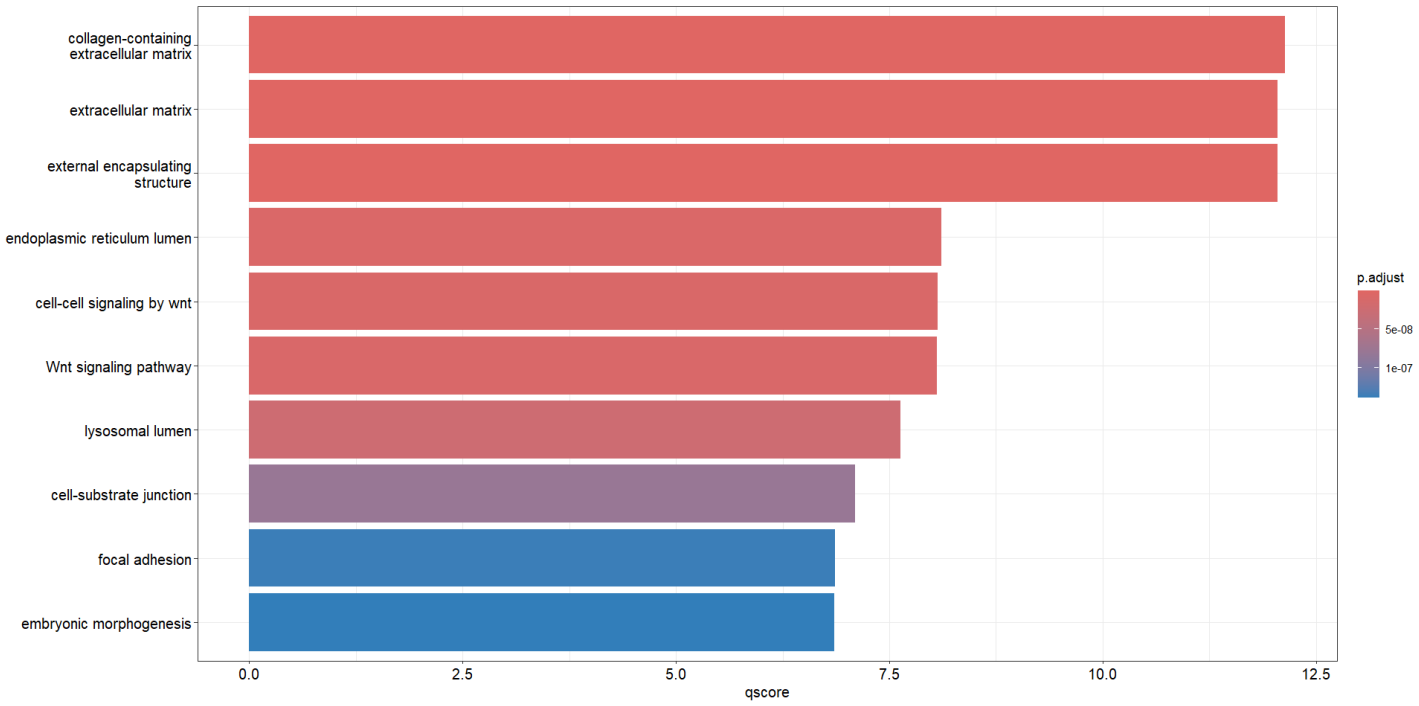


Figure 41. Top 10 upregulated pathways from GO enrichment analysis following differential expression in MUT 3.2.5.2 PGCs. This figure displays the top ten upregulated pathways identified through GO enrichment analysis after comparing MUT PGCs before (MUT-PG) and after dox-induced overexpression of MUT-NANOS1 (MUT-PGD). The pathways are ranked by their enrichment score (q-score), with the most enriched pathways presented at the top. X-axis (q-score) represents the enrichment score, a measure of the significance of pathway upregulation; higher q-scores indicate stronger upregulation. Y-axis lists the top ten upregulated pathways based on GO analysis, describing key biological processes or cellular components showing increased activity following MUT-NANOS1 overexpression. Color gradient: Each bar's color corresponds to the adjusted p-value (p.adjust), with red indicating stronger statistical significance and blue indicating relatively lower significance.

6. Altered RNA interactome size and preferred binding sites in WT- and MUT-NANOS1 cell lines

Following separate high-throughput analyses, the results for both WT-NANOS1 and MUT-NANOS1 cell lines were overlapped to identify components that were both bound by the respective NANOS1 protein and altered upon binding. **Fig.42** illustrates this overlap analysis for the pre-me and PGC stages of stem cell differentiation. Notably, the most enriched RNA interactome was observed in the MUT-NANOS1 cell line 3.2.5.2, where 149 targets were bound and altered at the pre-me stage (**Fig. 42**, top left), compared to only 19 targets in the WT-NANOS1 cell line 1.2.5.1 (**Fig. 42**, top right).

At the PGC stage, the enriched RNA interactome for the MUT-NANOS1 clone expanded to 385 targets (**Fig. 42**, bottom left), suggesting a more prominent role of MUT-NANOS1 during PGC specification than during the pre-me primitive-streak like stage. This observation is particularly significant given that the data were overlapped from the eCLIP-bound RNA interactome at the pre-me stage, which limits the precise identification of RNA targets bound specifically at the PGC stage. However, this approach allows for a temporal assessment of RNA targets that begin binding at the pre-me stage but have significant effects and play a prominent role during the PGC stage.

In contrast, the WT-NANOS1 protein showed a markedly reduced RNA interactome at the PGC stage, binding and altering only 6 targets (**Fig. 42**, bottom right). This minimal enrichment suggests that WT-NANOS1 exerts limited influence at this early point in human PGC development and may act later in the germ cell trajectory, particularly given that *in vitro* differentiation captures a stage before PGC migration and after primitive streak determination. Meanwhile, MUT-NANOS1 p.[(Pro34Thr; delSer78)] exerts a more pronounced impact, potentially disrupting pathways essential for PGC specification (**Fig. 41**), patients (Kusz-Zamelczyk et al., 2013).

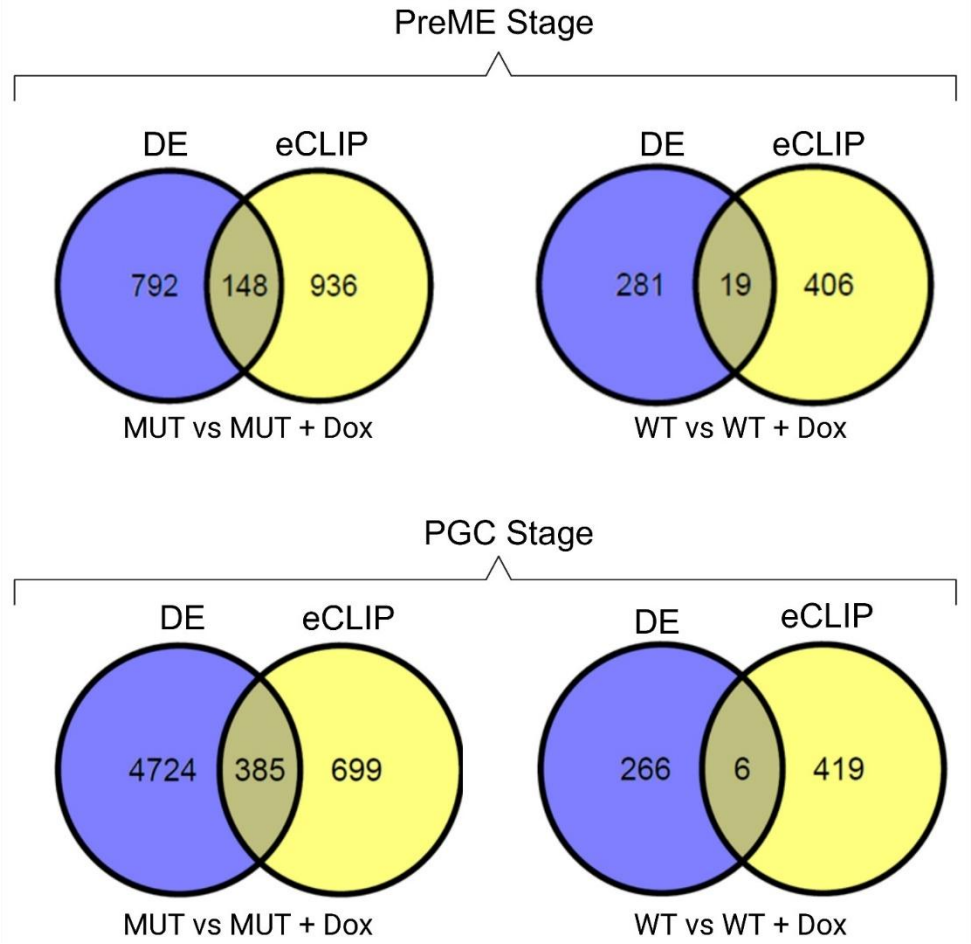


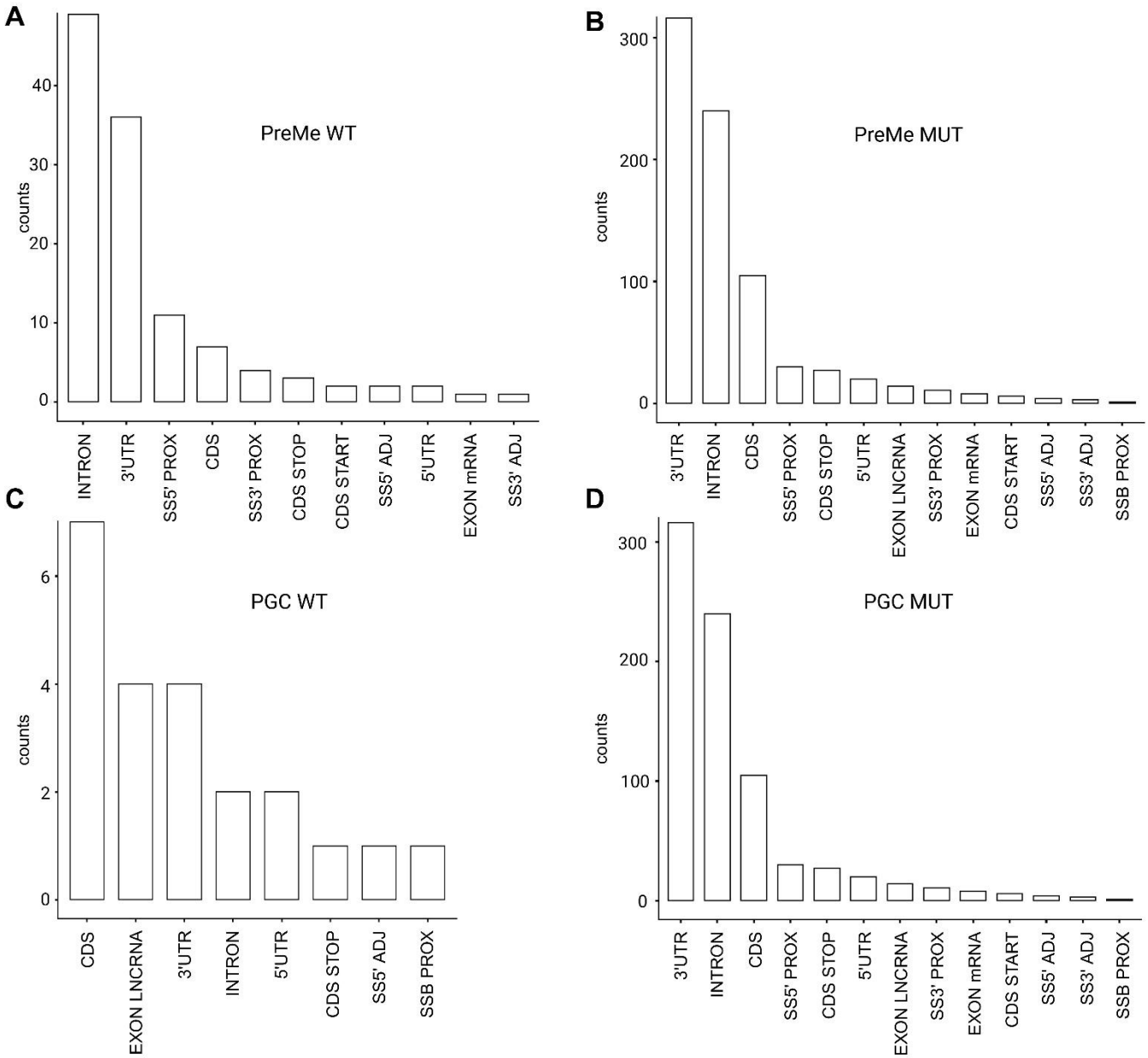
Figure 42. Venn diagrams show the overlap between differentially expressed (DE) genes and eCLIP-enriched genes at the pre-me and PGC stages. The data was derived from the MUT 3.2.5.2 cell line before and after dox induction at the pre-me (top left) and PGC (bottom left) differentiation stages. Similarly, results from the WT 1.2.5.1 cell line are shown before and after dox induction at the pre-me (top right) and PGC (bottom right) differentiation stages. These findings highlight the differential gene expression and binding sites associated with encoding genes representing MUT- and WT-NANOS1 conditions across developmental stages.

Furthermore, the enriched binding sites of the respective altered RNA interactome were analysed and visualised by integrating RNA sequencing and eCLIP data specifically from the pre-me and PGC stages (**Fig. 43A-D**). In the WT 1.2.5.1 cell line at the pre-me stage, introns are the predominant binding sites, followed by the 3'UTR (UTR3) (**Fig. 43A**). This pattern is consistent with previous enriched eCLIP data (**Fig. 31**) for

this cell line. By contrast, the 3'UTR becomes the dominant binding site in MUT-NANOS1, with a significant increase in binding counts compared to WT-NANOS1 (**Fig. 43A-B**), showing over 300 counts for MUT-NANOS1 versus approximately 35 counts for WT-NANOS1. Although introns still exhibit substantial binding for MUT-NANOS1, it occurs at a secondary level compared to the 3'UTR, with introns being enriched by over 200 counts, followed by the CDS with around 100 counts (**Fig. 43B**). These findings align with the broader and more diverse binding preferences for MUT-NANOS1 observed in the eCLIP-only interactome (**Fig. 32**), indicating a shift in RNA-binding specificity compared to WT-NANOS1.

At the PGC stage, the WT cell line exhibits limited enrichment, likely due to the very low overall bound interactome (**Fig. 43C** and **Fig. 42**, bottom right panel). Unlike in the pre-me stage (**Fig. 43A**), binding in the CDS dominates at the PGC stage, with over 6 counts (**Fig. 43C**), followed by similar counts in the exonic region of long non-coding RNA (EXON_LNCRNA) and the 3' UTR (approximately 4 counts each). Conversely, in the MUT 3.2.5.2 cell line at the PGC stage, the binding patterns are similar to those observed in pre-me (**Fig. 43B**), with the 3'UTR remaining the most enriched binding site (over 300 counts, **Fig. 43D**), followed by introns and CDS, both with lower counts compared to the 3'UTR. Overall, in both the pre-me and PGC stages, MUT-NANOS1 exhibits a shift in binding preferences, with a significantly increased enrichment in 3'UTR sites of target mRNAs and, to a lesser extent, introns and CDS (**Fig. 43B** and **43D**).

Lastly, to reflect the findings from the PGC stage (**Fig. 43C and 43D**), binding site densities for WT-NANOS1 and MUT-NANOS1 were normalized per kilobase against the total length of each region in the human genome. As shown in **Figure 43E**, WT-NANOS1 binding sites are also primarily enriched in the CDS followed by 3'UTR, with intronic binding sites being relatively sparse, likely due to the extensive length of intronic sequences. The binding profile of MUT-NANOS1 exhibits differences compared to WT, suggesting an altered RNA interactome. For the MUT-NANOS1 counterpart the 3'UTR and CDS are proportionally divided and dominate followed by 5'UTR (**Fig. 43E**). The marked 3'UTR enrichment, which was also the dominant feature in the bound and altered target mRNAs (**Fig. 43D**) for MUT-NANOS1, highlights a primary mechanism by which this protein may affect post-transcriptional regulation. Furthermore, a chi-square test ($\chi^2 = 21.029$, $p < 0.0001038$, **Fig. 43E**) confirms that these differences in binding site distribution between WT- and MUT-NANOS1 enriched RNA interactomes are statistically significant. Collectively, these data indicate that the p.[(Pro34Thr; delSer78)] variant significantly alters NANOS1's RNA binding preferences, potentially disrupting its normal regulatory functions.


A
PreMe WT
counts
INTRON
3'UTR
SS5' PROX
CDS
SS3' PROX
CDS STOP
CDS START
SS5' ADJ
5'UTR
EXON mRNA
SS3' ADJ
B
PreMe MUT
counts
3'UTR
INTRON
CDS
SS5' PROX
CDS STOP
5'UTR
EXON LNCRNA
SS3' PROX
EXON mRNA
CDS START
SS5' ADJ
SS3' ADJ
SSB PROX
C
PGC WT
counts
CDS
EXON LNCRNA
3'UTR
INTRON
5'UTR
CDS STOP
SS5' ADJ
SSB PROX
D
PGC MUT
counts
3'UTR
INTRON
CDS
SS5' PROX
CDS STOP
5'UTR
EXON LNCRNA
SS3' PROX
EXON mRNA
CDS START
SS5' ADJ
SS3' ADJ
SSB PROX

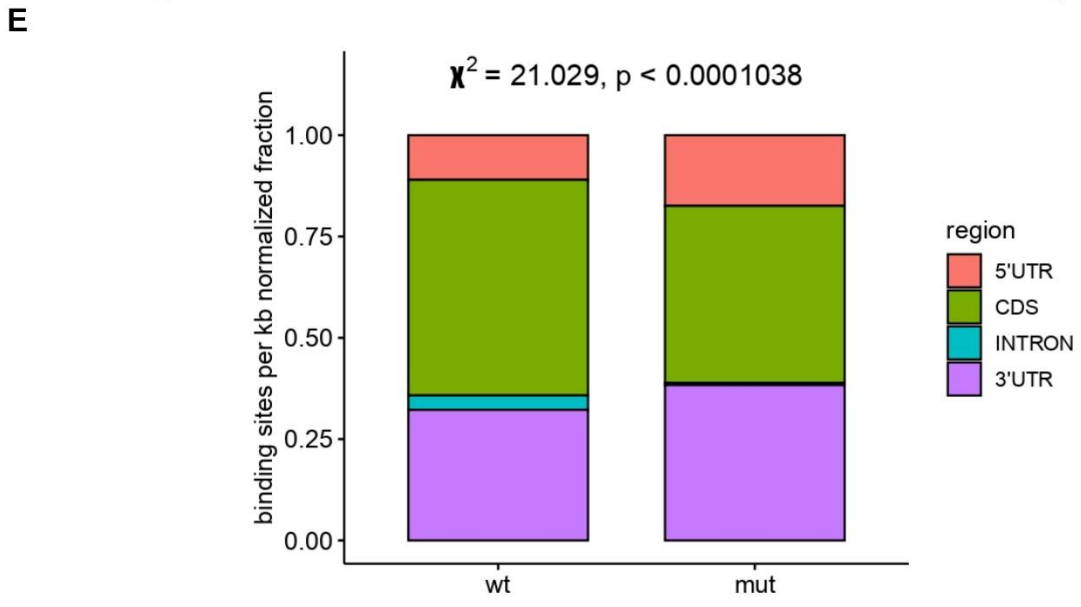

E
χ^2 = 21.029, p < 0.0001038
binding sites per kb normalized fraction
wt
mut
region
5'UTR
CDS
INTRON
3'UTR

Figure 43. Distribution of enriched binding sites from overlapping bound RNA interactome and altered RNA targets following eCLIP and bulk RNA sequencing, respectively. The data overlap focuses on WT- and MUT-NANOS1 cell lines across two developmental stages: pre-me and PGC. **A**- Distribution of binding site counts for the WT-NANOS1 1.2.5.1 cell line at the pre-me stage. The binding sites are predominantly found in introns followed by 3'UTR enriched sites. **B**- Binding RNA target/s site distribution in the MUT-NANOS1 3.2.5.2 cell line at the pre-me stage. There is a sharp increase in the count of binding sites in the 3'UTR regions, followed by introns, CDS, and other categories. **C**- Binding site counts for the WT-NANOS1 cell line at the PGC stage. Binding is most enriched in the CDS, exonic lncRNA regions, and 3'UTR. **D**- Distribution of binding site counts for the MUT-NANOS1 cell line at the PGC stage. The 3'UTR enriched sites retain the highest counts compared to other categories, which include introns, followed by CDS. **E**- Stacked bar plot showing the normalized fraction of binding sites per kilobase for WT- and MUT-NANOS1 target mRNAs across different transcript regions (5'UTR, CDS, intron, and 3'UTR). Binding sites were normalized to the total length of each respective region in the human genome. A chi-square test ($p < 0.0001038$) confirms that the distribution of binding sites across these regions differs significantly between WT- and MUT-NANOS1, indicating that the variant affects NANOS1's RNA binding preferences.

7. Enriched pathways of the altered and potentially bound RNA interactome of MUT-NANOS1 in pre-me cells

In light of the highly enriched and altered RNA interactome observed in the MUT-NANOS1 counterpart, the validation assays focused on elucidating its role in both the pre-me and PGC stages and the pathways affected. Gene set enrichment analysis (GSEA) revealed that the predominantly enriched pathways were downregulated to varying degrees (**Fig. 44**), primarily associated with cellular metabolism, macromolecule synthesis, and the cell cycle, specifically, regulation of primary metabolic processes, mitotic cell cycle, cell cycle phase transition, and macromolecule biosynthetic processes.

Most gene set intersections exhibited a median fold change between -0.75 and -1, indicating moderate downregulation of gene expression across these metabolic and cell cycle processes (**Fig. 44**). Some intersections displayed broader variability, suggesting that while certain genes within these pathways were significantly downregulated, others were less affected. For instance, the intersection labelled "regulation of primary metabolic process" showed consistent downregulation with a median fold change around -0.75. In contrast, more complex intersections involving multiple processes, such as those combining cell cycle and metabolic processes, demonstrated a wider range of variability, indicating heterogeneity in gene responses within these pathways (**Fig. 44**).

The downregulation of "mitotic cell cycle" and "cell cycle phase transition" suggests a potential reduction in cell division rates and alterations in cell cycle duration. Concurrent suppression of "macromolecule biosynthetic processes" and "nucleobase-containing compound metabolic processes" implies an impact on RNA, DNA, and protein synthesis, potentially hindering cellular growth and function.

In vivo studies have shown that only a fraction of epiblast cells commit to the PGC lineage in mouse and pig embryos, highlighting a significant cell-intrinsic barrier for PGC fate determination (Tang et al., 2022). Similarly, *in vitro* differentiation of human PGCLCs results in only 10–40% of pre-me cells successfully generating PGCs, underscoring the necessity for a precise epigenetic state, appropriate ME transcription factor dosage (Tang et al., 2022), and specific cell cycle stages (Pauklin and Vallier, 2013).

Therefore, MUT-NANOS1 appears to interfere with the metabolic and proliferative requirements for PGC formation *in vitro*, as evidenced by the widespread downregulation of key pathways (**Fig. 44**). This disruption signifies a substantial impairment of the cellular machinery necessary for these cells to progress in their differentiation, potentially impairing PGC development by interfering with essential biosynthetic and proliferative functions.

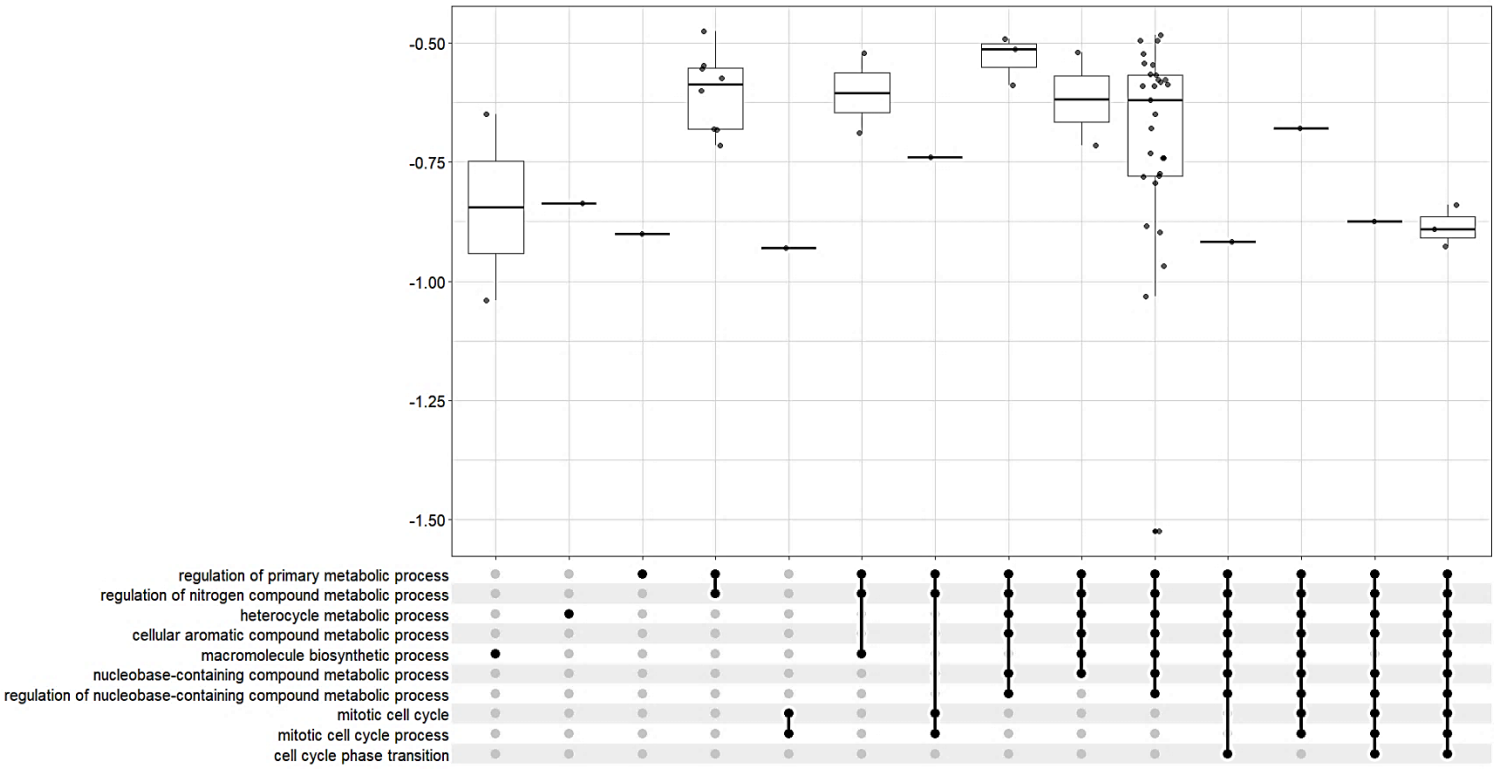


Figure 44. UpSet plot illustrating the results of GSEA for overlapped genes enriched from eCLIP and RNA-seq data in the MUT-NANOS1 cell line at the pre-me stage. Each row denotes a distinct biological process enriched in the dataset, with intersections represented below to show specific combinations of gene sets associated with each process. The y-axis depicts normalized enrichment scores, with higher values indicating stronger enrichment within each biological pathway. Notable enriched processes include "regulation of primary metabolic processes," "cell cycle phase transition," and "nucleobase-containing compound metabolic processes," suggesting key regulatory pathways influenced by MUT-NANOS1 at the pre-me differentiation stage.

In further examining the downregulated cell cycle processes (**Fig. 44**), enrichment plots from the GSEA were generated to confirm their repression (**Fig. 45**). The progression of the cell cycle is fundamentally linked to the self-renewal and differentiation of stem cells, enabling tissue specification (Pauklin and Vallier, 2013). In both graphs presented in **Fig. 45**, the enrichment scores for the "cell cycle phase transition" (top) and "mitotic cell cycle" (bottom) gene sets trend negatively, as indicated by the declining green lines. This suggests that most genes associated with these processes are downregulated, with a concentration toward the left side of the ranked list (**Fig. 45**).

The observed downregulation of cell cycle-related genes in pre-me cells overexpressing MUT-NANOS1 implies a broader disruption of cell cycle-dependent differentiation signals. For instance, Pauklin and Vallier (2013) demonstrated that hESC differentiation is tightly governed during the G1 phase of the cell cycle, primarily by the Activin/Nodal pathway in coordination with cyclin D proteins. Their findings highlight Smad2/3 phosphorylation as a critical 'gating' mechanism for directing cells toward specific lineages. The presence of MUT-NANOS1 may disrupt both the cell cycle and the proper modulation of the Activin/Nodal pathway, particularly affecting Smad2/3 phosphorylation and its downstream signaling. Consequently, these disruptions could alter the normal G1-based regulation of differentiation potential. Since PGC specification in early development requires tight coordination of cell cycle progression and signalling pathways, any disruption of cell cycle regulation due to MUT-NANOS1, as highlighted in **Fig. 45**, may prevent cells from entering the appropriate differentiation window.

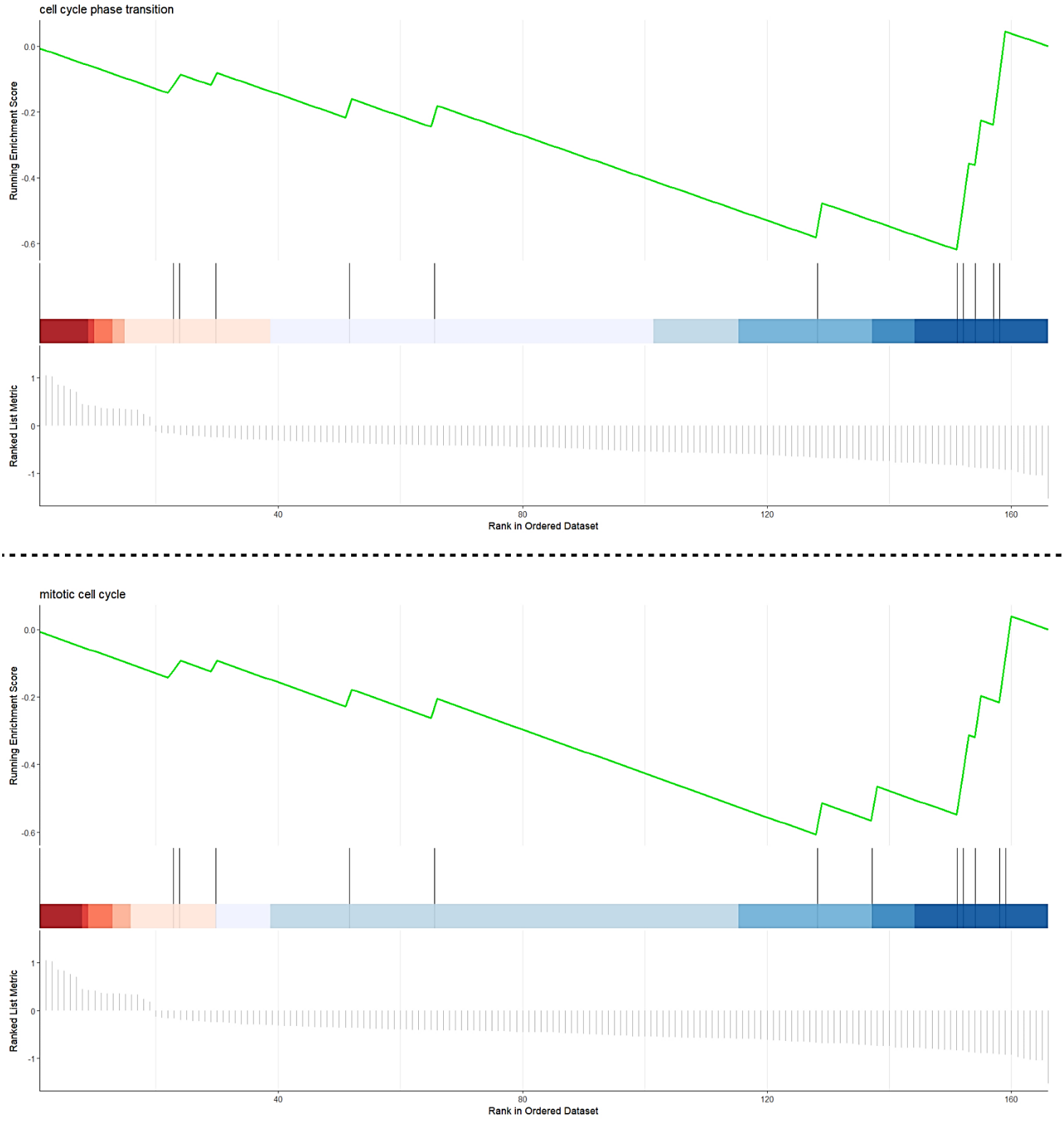


Figure 45. Enrichment plots from GSEA of genes bound and differentially expressed following overexpression of MUT-NANOS1 at pre-me. The top plot illustrates enrichment for the "cell cycle phase transition" gene set, while the bottom plot shows enrichment for the "mitotic cell cycle" gene set. In each plot, the green line represents the running enrichment score (ES) across the ranked gene list, with peaks indicating the ES for each gene set. Vertical black lines denote the positions of genes from each gene set within the ranked list, organized by their correlation with MUT-

NANOS1 overexpression. The color gradient below each plot indicates regulation direction, with red representing upregulated genes and blue representing downregulated genes.

Following the initial GSEA results, which revealed extensive downregulation across key pathways (**Fig. 44**), a focused GO analysis was conducted on the downregulated genes bound by MUT-NANOS1 in pre-me cells (**Fig. 46**). The GO plot highlights widespread downregulation of critical developmental pathways (**Fig. 46**), such as tissue morphogenesis, mesenchyme development, and epithelial morphogenesis, indicating that MUT-NANOS1 expression broadly suppresses essential developmental and differentiation processes required for cellular structural and functional organization.

Downregulation of pathways crucial for stem cell population maintenance and growth factor responses (**Fig. 46**) further indicates impaired self-renewal and diminished responsiveness to external signals, both essential features of proliferative PGC precursors. Additionally, suppressed mesenchyme and heart development pathways (**Fig. 46**) point to an overarching inhibition of early differentiation, potentially restricting progress toward the PGC fate.

Notably, the GO analysis highlights a crucial decrease in WNT signalling in pre-me cells overexpressing MUT-NANOS1 (**Fig. 46**), an especially significant result given WNT pathway's well-established role in embryonic development and germ cell lineage specification. By potentially binding and suppressing key WNT pathway components in pre-me, MUT-NANOS1 likely hampers PGC development and impairs the wider differentiation potential of these cells. Overall, the suppression of essential pathways such as cellular morphogenesis, stem cell maintenance, and differentiation signaling (**Fig. 46**) underscores how MUT-NANOS1 disrupts fundamental biological processes in pre-me differentiation stage.

Overall, these results parallel previous GO analyses based solely on RNA sequencing of MUT-NANOS1 PGCs (**Fig. 39**), reinforcing the inhibitory role of MUT-NANOS1 on processes critical for early PGC development. This interference likely impedes the developmental and signaling programs needed for pre-me cells to transition toward PGCs.

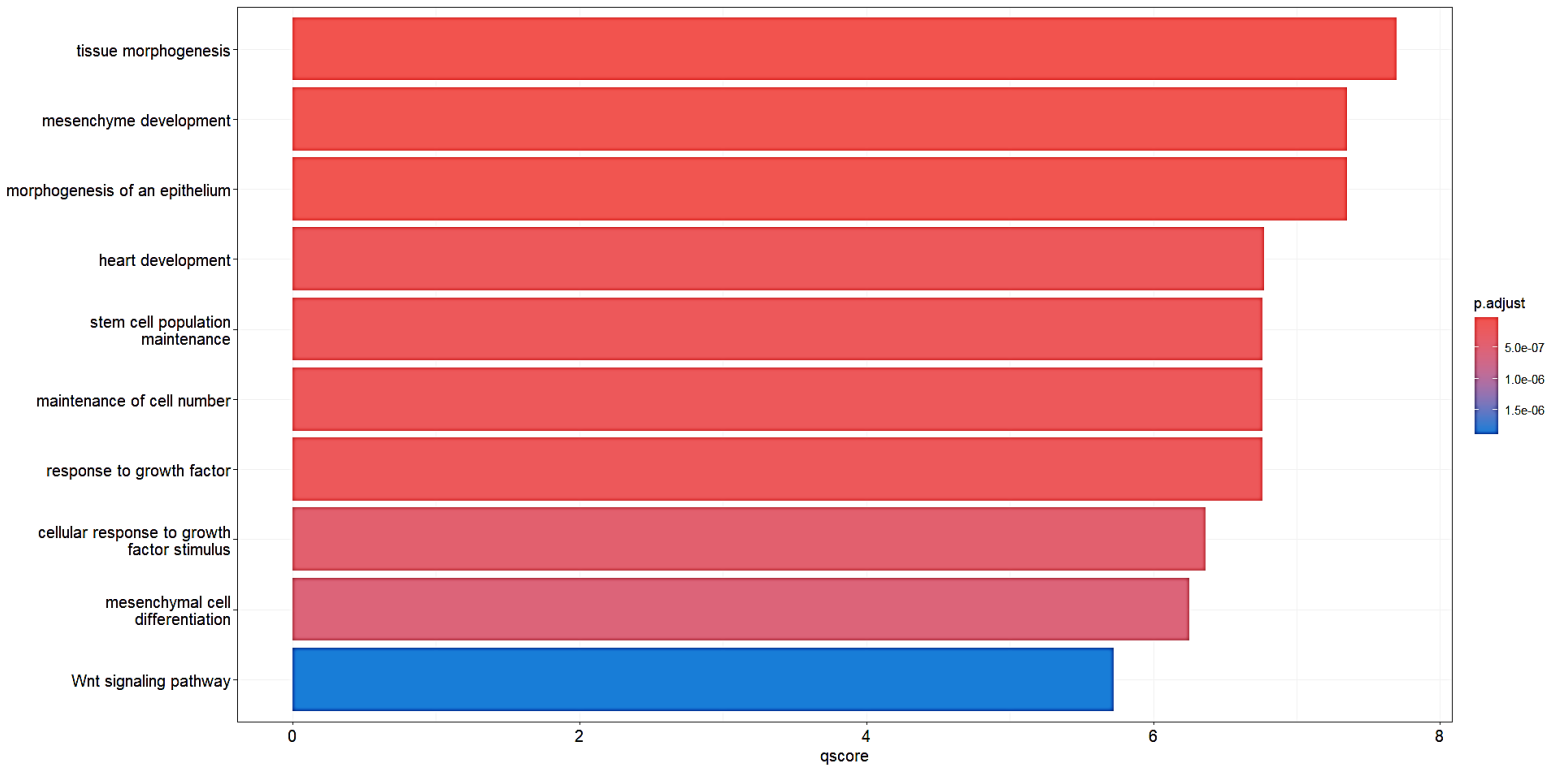


Figure 46. Gene ontology (GO) analysis depicting the top 10 downregulated pathways derived from mRNAs that are both differentially expressed and bound by MUT-NANOS1 at the pre-me stage. The bar chart ranks pathways by q-score (p.adjust), with tissue morphogenesis, mesenchyme development, and morphogenesis of an epithelium among the most significantly downregulated processes. The color gradient represents the adjusted p-values (q-score), with darker red indicating higher statistical significance (p.adjust value cutoff < 0.05).

8. Several WNT pathway components were altered by MUT-NANOS1 at the pre-me stage and, to an extent, at the PGC stage

An in-depth exploration of components within the WNT signalling pathway was undertaken for further validation of their roles. To elucidate the interconnections among these components, a cnetplot was generated following GO analysis (**Fig. 46** and **Fig. 47**). In this plot, the size of each node corresponds to its significance or degree of association within the WNT signalling network, with larger nodes indicating higher relevance. Terms such as "canonical WNT signalling pathway," "regulation of WNT signalling pathway," and "cell-cell signalling by WNT" represent key pathway elements negatively affected following overexpression of MUT-NANOS1 (**Fig. 47**). Notably, the highlighted components are directly bound by MUT-NANOS1 as per eCLIP data, suggesting direct regulation via the MUT-NANOS1 protein.

8.1 WNT pathway components potentially bound and altered by MUT-NANOS1 at the pre-me differentiation stage

Several mRNAs, including CD24, CCND1, FZD10, SFRP2, and TCF7L1, exhibit marked downregulation and likely represent key targets of MUT-NANOS1 overexpression and potential direct binding. According to

Teague et al. (2023), in an *in vitro* model of early human embryonic patterning, a BMP-WNT-Nodal hierarchy establishes dynamic gradients crucial for PGCLC differentiation, with the pre-me model first inducing a primitive streak–like state before exogenous BMP signalling.

The majority of downregulated WNT components depicted in **Fig. 47** belong to the canonical pathway, known for its pivotal roles in cell fate determination, proliferation, and survival (Mohammed et al., 2016). This is consistent with the findings of the affected pathways from the GSEA and GO analysis (**Fig. 45** and **Fig. 46**).

Notably, several of the most downregulated mRNAs in the cnetplot (**Fig. 47**), such as FZD10, influence, regulate and/or are canonical WNT targets. Specifically, FZD10 encodes a member of the Frizzled (FZD) receptor family, which is responsible for activating the canonical WNT pathway. This activation occurs when WNT ligands bind to the extracellular cysteine-rich domain of the FZD receptor and to the co-receptor LRP5/6 (Mohammed et al., 2016).

CCND1, another prominently downregulated WNT target (**Fig. 47**), is recognized by TCF/LEF transcription factors and encodes cyclin D1, essential for the G1-to-S phase transition. *In vitro* studies show that degradation-resistant β-catenin mutants significantly increase cyclin D1 expression and promoter activity (Shtutman et al., 1999). eCLIP data (not shown) suggest MUT-NANOS1 may bind the cyclin D1 3′UTR, affecting its stability. This downregulation correlates with reduced cell-cycle progression observed in GSEA enrichment plots (**Fig. 45**) and the UpSet plot (**Fig. 44**).

CD24 is also significantly downregulated (**Fig. 47**). This WNT target gene is commonly used as a surface marker for identifying putative cancer stem cell populations due to the critical role of WNT signalling not only in stem cell proliferation but also in oncogenesis (Wright et al., 2008). Overall, CD24 mRNA was identified as primed state marker (Collier et al., 2017) enriched in the surface of post-implantation epiblast cells. Its presence at the pre-me stage aligns with the notion that PGC progenitors exhibit primed pluripotency or an intermediate state between naïve and primed pluripotency leaning toward the latter (Chen et al., 2019).

In addition to target mRNAs, the WNT modulator SFRP2 is among the mRNAs to be significantly downregulated (**Fig. 47**). Notably, it plays an important role in regulating the WNT pathway and mediating PGC competence in certain pre-me progenitor subpopulations (Castillo-Venzor et al., 2023). Typically, SFRP2 antagonizes WNT by competing with WNT morphogens for FZD binding, although its role can be context-depended. For example, in primary cardiac fibroblasts, SFRP2 enhanced canonical WNT/β-catenin signaling by promoting nuclear β-catenin accumulation, and thereby driving proliferation, glucose-based metabolism, and ECM remodeling (Lin et al., 2016).

TCF7L1, a WNT pathway component known to primarily repress or weakly activate WNT target genes, is among the transcripts to be highly significantly downregulated (**Fig. 47**). This factor is typically the most highly expressed TCF/LEF family member in undifferentiated primed hESCs, where it binds and suppresses genes crucial for gastrulation and primitive streak formation, including NODAL, WNT3, and BMP4, thereby maintaining pluripotency (Sierra et al., 2018). Under normal conditions, exogenous BMP4

downregulates TCF7L1 ahead of PGCLC differentiation to allow ME induction (Sierra et al., 2018). As the pre-me stage shown in **Fig. 47** reflects cells before BMP4 treatment, the premature downregulation of TCF7L1 suggests a disruption of its repressive function in WNT signaling and pluripotency.

The downregulation of the non-canonical WNT pathway is also represented in **Fig. 47** but to a lesser extent, highlighted by non-canonical WNT members such as ROR1. This pathway is more often associated with differentiation, cell polarity, and migration (Mohammed et al., 2016).

Overall, these collective changes suggest MUT-NANOS1-mediated binding and downregulation of crucial WNT signaling components, repressing cell-cycle progression and differentiation cues. Since WNT activity underpins the primitive streak–like induction and Nodal activation essential for the pre-me to PGCLC transition, this repression likely impedes proper PGC differentiation.

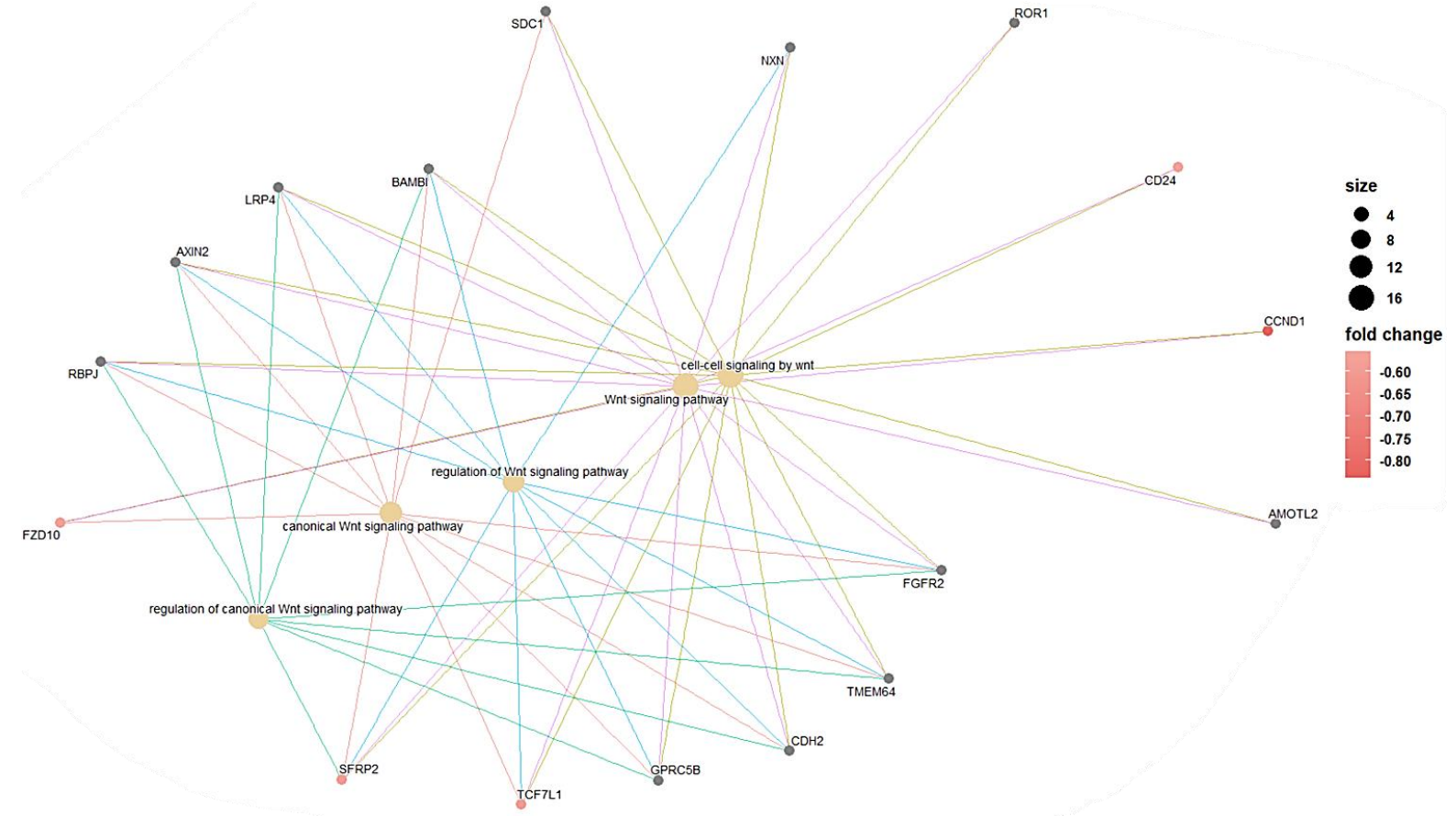


Figure 47. Cnetplot illustrating the downregulated WNT signaling pathway within the enriched RNA interactome of MUT-NANOS1 following its overexpression at the pre-me stage. This network visualization highlights genes involved in the WNT signaling pathway, with nodes representing individual genes and pathway terms. The size of each gene node corresponds to the degree of connectivity, indicating its involvement across multiple pathways. Node color reflects the fold change in expression, with deeper red indicating a greater degree of downregulation (fold change). Key pathways, such as "canonical Wnt signaling pathway," "regulation of Wnt signaling pathway," and "cell-cell signaling by Wnt," are central, emphasizing the suppression of Wnt signaling-related processes in response to MUT-NANOS1 overexpression.

To validate the RNA sequencing findings, real-time quantitative PCR (RT-qPCR) was conducted on two independently generated MUT-NANOS1 cell lines, 3.1.8 and 3.2.5.2, confirming significant downregulation of WNT pathway-related targets in both lines (**Fig. 48**). This consistency implies that the observed effects stem from the MUT-NANOS1 protein rather than cell line-specific variation.

Although all tested WNT components showed reduced expression (**Fig. 48**), not all were directly bound by MUT- NANOS1. Overall, binding by MUT-NANOS1 is based on eCLIP results however, further assays (e.g. luciferase) would be needed to fully validate these binding events. Nevertheless, transcripts such as TCF4, TCF7L1, FZD10, SFRP2, BAMBI, and SDC1, were both bound and downregulated (**Fig. 47** and **Fig. 48**), implicating a mechanism by which MUT-NANOS1 represses WNT signalling.

Other pivotal WNT regulators, including LEF1, TCF7, DACT1, and WNT5B, were also significantly downregulated (**Fig. 48**), likely as secondary effects rather than through direct binding by MUT-NANOS1.

DACT1, for example, although often regarded as a WNT antagonist, exhibits context-dependent roles. In human placenta development, Hou et al. (2015) demonstrated that increased DACT1 mRNA correlates with enhanced WNT signalling, including elevated expression of target genes such as cyclin D1 (notably downregulated in **Fig. 47**), cyclin D2, FOSL1, and JUN. Its reduced expression here underscores the broader impact of WNT modulation on directing cells toward a primitive streak–like and PGC fate.

WNT signalling pathway relies on TCF/LEF transcription factors that serve as binding partners for β-catenin, the primary mediator of WNT-mediated transcription (Cadigan and Waterman, 2012). When WNT ligands, such as the downregulated WNT5B highlighted in **Fig. 48**, and shown by Castillo-Venzor et al. (2023) to interact with SFRP2 in PGC-competent pre-me subpopulations, bind to their receptors (such as FZD10, also downregulated as seen in **Fig. 47**), β-catenin stabilizes and moves into the nucleus. Inside the nucleus, it associates with TCF/LEF factors attached to WNT-response elements (WREs) on the DNA, which then activates WNT target genes. Each TCF/LEF factor has a distinct regulatory capacity, with TCF7 and LEF1 functioning as strong activators (Cadigan and Waterman, 2012). Consequently, the downregulation of these factors, though not directly bound by MUT-NANOS1, emphasizes their critical role in positively modulating WNT signalling and steering differentiation toward the primitive streak and PGCs.

Among directly bound targets, transcription factor 4 (TCF4) mRNA (**Fig. 48**) is a central nuclear effector in the canonical WNT pathway, known for its dual functions in transcriptional activation and repression (Salinas et al., 2008). In the absence of WNT signals, TCF4 binds to WNT target gene promoters to inhibit their expression. Upon WNT pathway activation, β-catenin stabilizes in the cytoplasm, translocates into the nucleus, and forms a complex with TCF4, thereby activating crucial WNT target genes (Lee et al., 1999). MUT-NANOS1–mediated TCF4 downregulation in both MUT cell lines (**Fig. 48**) may hinder its normal regulation of WNT-dependent gene expression, further impeding pre-me cells' differentiation potential.

TCF4 expression is additionally regulated by microRNAs, which can modulate WNT signalling and cell proliferation (Ma et al., 2024). Inhibiting miR-140-y rescues TCF4 activity, thus enhancing WNT-related

gene expression and cell growth. Consequently, reduced TCF4 expression in pre-me cells, whether via MUT-NANOS1 binding and mediated repression or *via* miRNA influence, may impede WNT pathway activity and compromise differentiation toward the primitive streak or PGC lineage.

Lastly, BAMBI and Syndecan-1 (SDC1), validated for their downregulation (**Fig. 48**) and bound through 3'UTR and CDS by MUT-NANOS1, are not core components of the WNT pathway but participate in critical crosstalk with other signalling pathways. BMP and Activin Membrane-Bound Inhibitor (BAMBI), a negative TGF-β regulator, is capable of interfering with the pathway's usual role in growth arrest often driven by β-catenin activation, especially in cancer stem cells (Sekiya et al., 2004). In developmental contexts, low BAMBI expression fosters PGC competence by allowing active Nodal and WNT signaling (Castillo-Venzor et al., 2023); however, subsequent downregulation is crucial for driving pre-me cells toward PGC fate. Consequently, the binding of MUT-NANOS1 to BAMBI inducing premature downregulation of the transcript may disrupt this balance, impairing germ cell development.

SDC1 similarly plays a key role in modulating the WNT/β-catenin pathway as its silencing diminishes β-catenin levels in neural progenitor cells, reducing cell growth (Wang et al., 2012), while soluble SDC1 can activate WNT1, prompting mammary epithelial tumour formation (Liao et al., 2020). Downregulation of both BAMBI and SDC1 (**Fig. 48**) underscores the broader network of developmental signals disrupted by MUT-NANOS1; this premature repression leads to aberrant WNT regulation and ultimately hinders the normal transition from pre-me to PGC.

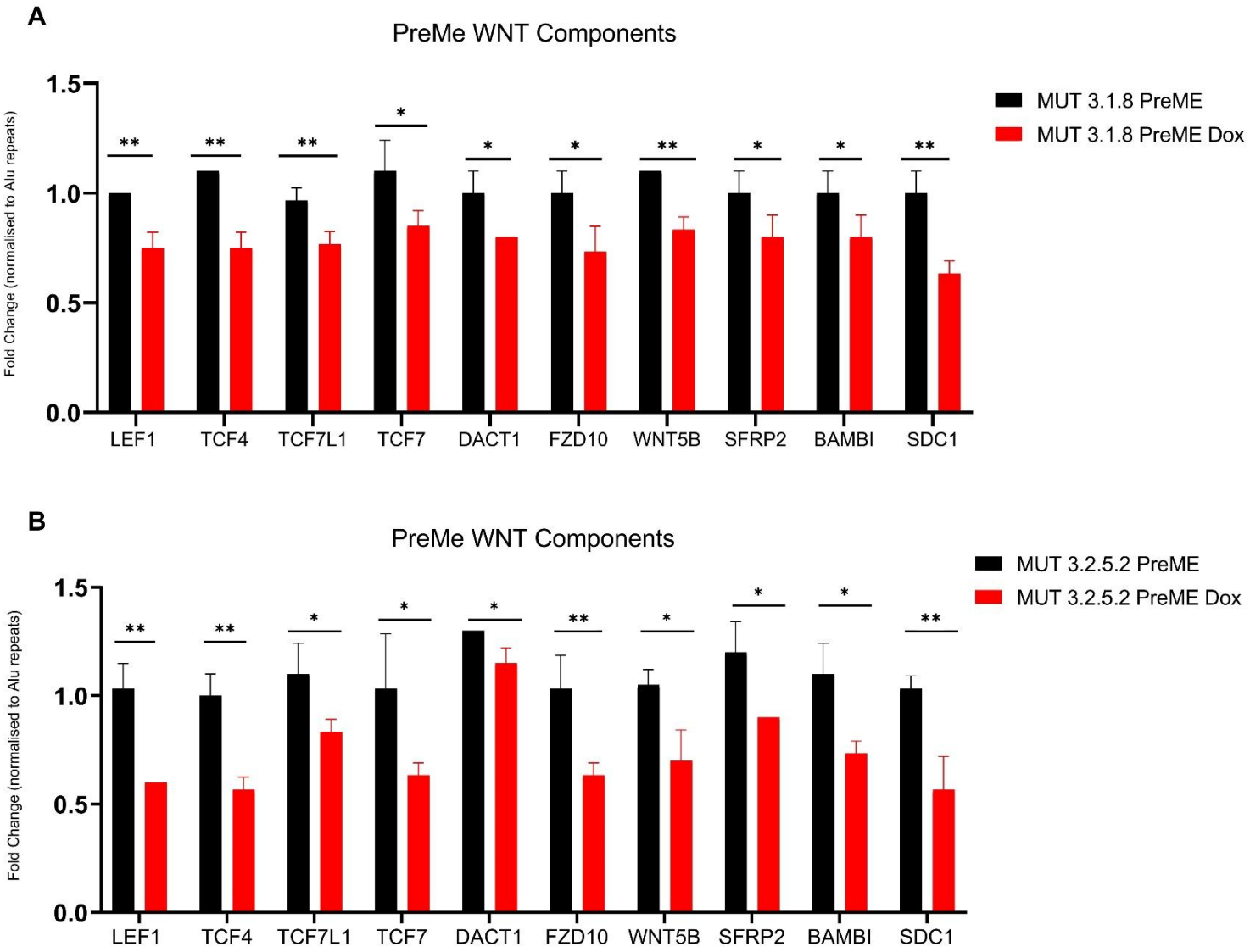


Figure 48. RT-qPCR analysis of WNT pathway components in pre-me stage of MUT-NANOS1 cell lines. **A**- and **B**- The bar plots represent fold changes normalized to Alu repeats, comparing untreated (black bars) and dox-treated (pink bars) biological replicates (3 per sample tested). Panels **A** and **B** show expression data for two different MUT-NANOS1 cell lines: 3.1.8 and 3.2.5.2, respectively. Statistical significance was assessed using multiple unpaired one-tailed t-tests, with $p < 0.05$. Significantly downregulated WNT pathway components include: Lymphoid Enhancer Binding Factor 1 (LEF1); Transcription Factor 4 (TCF4); Transcription Factor 7 Like 1 (TCF7L1); Transcription Factor 7 (TCF7); Dapper Antagonist of Catenin 1 (DACT1); Frizzled Class Receptor 10 (FZD10); Wnt Family Member 5B (WNT5B); Secreted Frizzled Related Protein 2 (SFRP2); BMP And Activin Membrane Bound Inhibitor (BAMBI) and Syndecan 1 (SDC1). Asterisks indicate levels of significance: *$p < 0.05$ and **$p < 0.01$. Standard deviation bars are plotted, and outlier replicate/s were removed accordingly.

8.1.1 RT-qPCR validation of WNT pathway targets AXIN2, CCND1, CD24, and MYC in MUT-NANOS1 cell lines during pre-me induction

In addition to confirming the downregulation of WNT pathway components in both MUT-NANOS1 cell lines (**Fig. 48**), RT-qPCR validated reduced expression of WNT target genes cyclin D1, CD24, AXIN2, and c-MYC (MYC) in MUT cell lines 3.1.8 and 3.2.5.2 (**Fig. 49A-D**). This downregulation can be partly attributed to the suppressed WNT pathway, as indicated by reduced expression of its components (**Fig. 48A-B**), but it

is also driven by MUT-NANOS1 binding since these target mRNAs were enriched in the MUT-NANOS1 eCLIP data (data not shown).

Among the WNT targets, cyclin D1 and CD24 are central to cell cycle progression and post-implantation epiblast primed pluripotency, respectively, underscoring their importance for proliferation and lineage commitment upon WNT and Nodal pathway activation. Additionally, AXIN2 and MYC are well-established targets of WNT/β-catenin signaling (Rennoll et al., 2015). AXIN2 operates as part of the destruction complex, providing a negative feedback mechanism to limit WNT activity (Leung et al., 2002). Although initiated by canonical WNT signals, AXIN2 mRNA may be prematurely degraded when bound at its 3′UTR by MUT-NANOS1.

Overall, AXIN2 displays dynamic, temporally and spatially restricted expression during embryonic development, including in the primitive streak of chick and mouse embryos (Eckei et al., 2016; Huelsken et al., 2000), underscoring its role in WNT-dependent processes. In human embryonic stem cells, AXIN2 stabilizes cytoplasmic β-catenin to maintain pluripotency, while nuclear β-catenin drives differentiation (Kim et al., 2013). In this study, the downregulation of both cytosolic and nuclear WNT pathway factors disrupt pluripotency and differentiation.

MYC, on the other hand, is a TF that drives cellular proliferation by activating target genes (Dang et al., 2006). When nuclear β-catenin increases, it induces genes such as cyclin D1 and MYC, influencing critical cell cycle phases (Lecarpentier et al., 2019). MYC also represses differentiation-associated genes during induced pluripotent stem cell (iPSC) formation, although it does not necessarily maintain pluripotency factor expression (Wey and Knoepfler, 2010). Instead, it supports the high metabolic demands of pluripotent cells (van Riggelen et al., 2010; Dominguez-Sola et al., 2008). Notably, regulation of primary associated metabolic pathways was downregulated following MUT-NANOS1 overexpression in cell line MUT 3.2.5.2 (**Fig. 44**, UpSet plot). Furthermore, MYC can activate or repress specific genes, regulate miRNAs and rRNAs, drive DNA synthesis, and affect chromatin structure, in part by relieving transcriptional pausing via P-TEFb (Wey and Knoepfler, 2010; Rahl et al., 2010). Its broad regulatory functions may arise from widespread euchromatic binding linked to histone modifications such as H3K9 acetylation and H3K4 methylation (Martinato et al., 2008). Therefore, the AXIN2–MYC axis is central to cell proliferation and development, and by repressing these key transcripts, MUT-NANOS1 disrupts critical WNT-driven processes, ultimately impairing pluripotency maintenance, cell cycle progression, and normal differentiation toward the primitive streak–like and PGCLC fates.

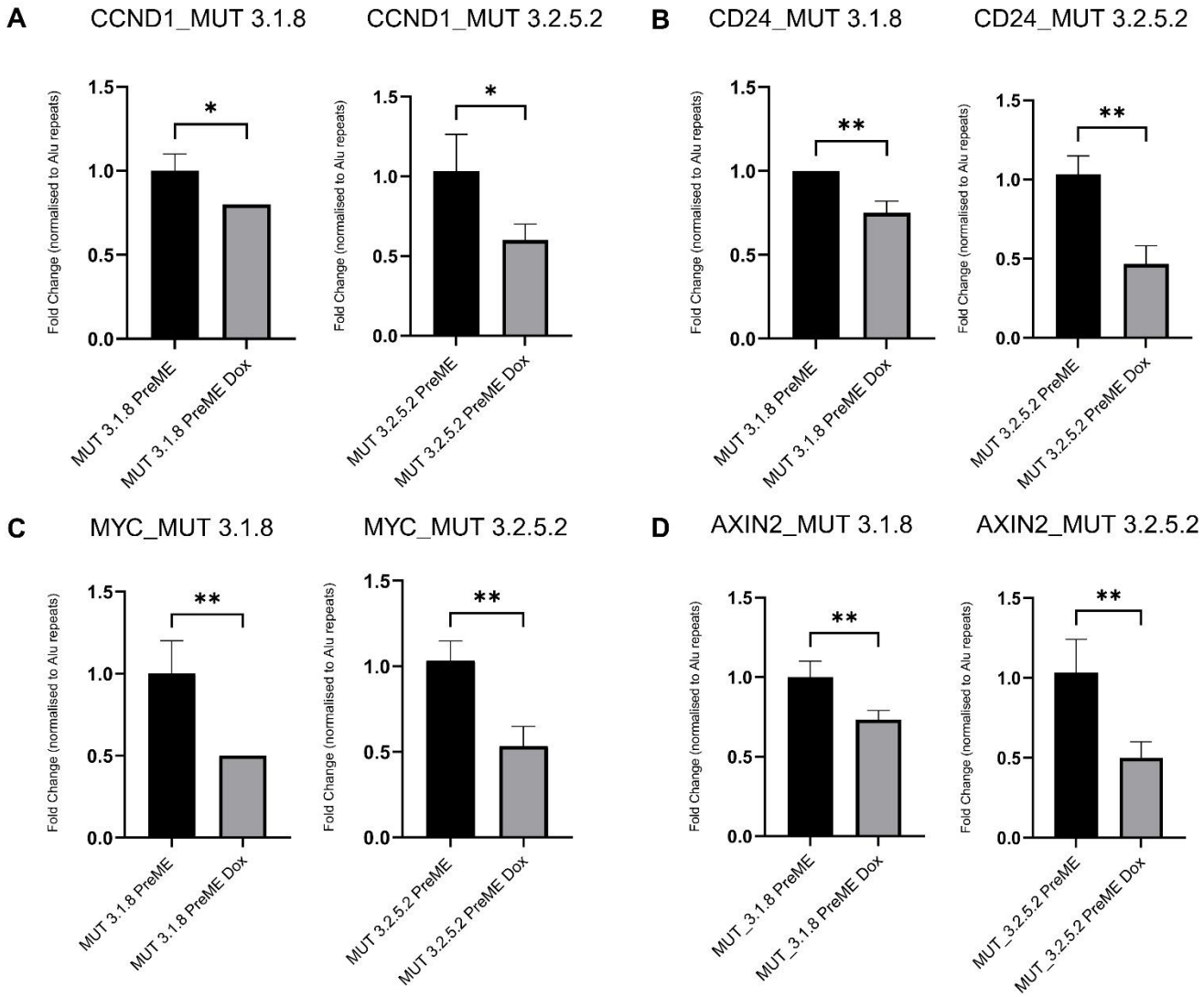


Figure 49. RT-qPCR analysis of select WNT target genes in MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 at the pre-me stage. The bar plots represent fold changes normalized to Alu repeats, comparing untreated (black bars) and dox-treated (grey bars) biological replicates (three per condition). Each panel shows a specific WNT target mRNA, and statistical significance was assessed using unpaired one-tailed t-tests, with significance applied when $p < 0.05$. The following WNT target mRNAs were significantly downregulated upon dox treatment in both MUT-NANOS1 3.1.8 and 3.2.5.2 cell lines: **A**- CCND1 (cyclin D1); **B**- Cluster of Differentiation 24 (CD24); **C**- MYC Proto-Oncogene, bHLH TF (MYC); **D**- Axis Inhibition Protein 2 (AXIN2). Asterisks indicate levels of significance: *$p < 0.05$ and **$p < 0.01$. Standard deviation bars are plotted per sample with outlier replicate/s removed accordingly.

8.1.2 MUT-NANOS1 binding to MYC, SFRP2, and CDH2 mRNAs may be mediated via the SAAGAA motif in pre-me cells

In addition to confirming downregulated targets in pre-me cells, motif analysis was performed on the overlapping bound mRNAs to determine whether MUT- or WT-NANOS1 protein recognizes specific RNA motifs. Although most motifs appeared at low frequencies or were found in only one RNA target, the SAAGAA (GAAGAA) motif was identified in 16.87% of the assessed sequences and shared among three validated targets, MYC, SFRP2, and CDH2, specifically within their CDS (**Fig. 50A**). This finding suggests

a mechanism by which MUT-NANOS1, rather than WT-NANOS1, might specifically bind to these transcripts carrying this common motif, all of which are interconnected with the WNT pathway, although this would require further validation via luciferase assays.

The absence of the motif and corresponding targets in WT-NANOS1 analyses (data not shown) indicates a gain-of-function in the MUT-NANOS1 p.[Pro34Thr; Ser78del] variant, enabling novel binding interactions that disturb WNT pathway regulation.

While MYC has been characterized as a WNT target gene and SFRP2 as a WNT pathway modulator, CDH2 mRNA stands out as another key factor tied to this pathway from a different perspective. Time-course analysis indicates that in MUT-NANOS1 cell line, CDH2 expression reaches its highest levels during the pre-me and somatic (SO) stages following PGC differentiation (**Fig. 50B** and **Fig. 50F-G**). However, in the control cell line W15, before and after dox addition (**Fig. 50F-G**), CDH2 expression increases steadily from the stem cell (SC) stage through pre-me and PGCs (PG), peaking at SO. In contrast, the MUT-NANOS1 cell line shows a decline in CDH2 expression from pre-me to PGC stage, especially when comparing against its own control before dox induction (**Fig. 50B**). This indicates that the downregulation of CDH2 mRNA is observed not only at the pre-me stage, validated for its downregulation in MUT cell lines (**Fig. 50B and Fig. 50H**), but also throughout the transition from pre-me to PGC, implying that MUT-NANOS1 binding affects CDH2 mRNA expression consistently throughout differentiation. Furthermore, the statistically significant downregulation of CDH2 mRNA compared to the control cell line W15 (**Fig. 50H**) was also evident when validating its levels within both MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 before and after dox induction (**Figs. 50F-G**). Although there may be some influence from the cell line itself, MUT-NANOS1 overexpression notably intensifies the downregulation of CDH2 at the pre-me stage, with normalized counts decreasing from around 600 before dox induction to roughly 400 afterward (**Fig. 50F-G**).

During gastrulation-related epithelial-mesenchymal transition (EMT), nascent mesoderm cells replace CDH1 with CDH2 and upregulate T expression, creating a posterior strip of CDH2/T-expressing cells by E6.5 in mice models (Hernández-Martínez et al., 2019). This switch from CDH1 (E-cadherin) to CDH2 (N-cadherin) leads to loss of cell–cell junctions, increased cell motility, and reorganized adherens junctions involving β-catenin (Wal and van Amerongen, 2020; Oda and Takeichi, 2011; Zhang et al., 2009). Temporarily elevated levels of free β-catenin may activate WNT signalling, promoting partial progression (10–40% of cells) toward PGCLCs during BMP4 exposure.

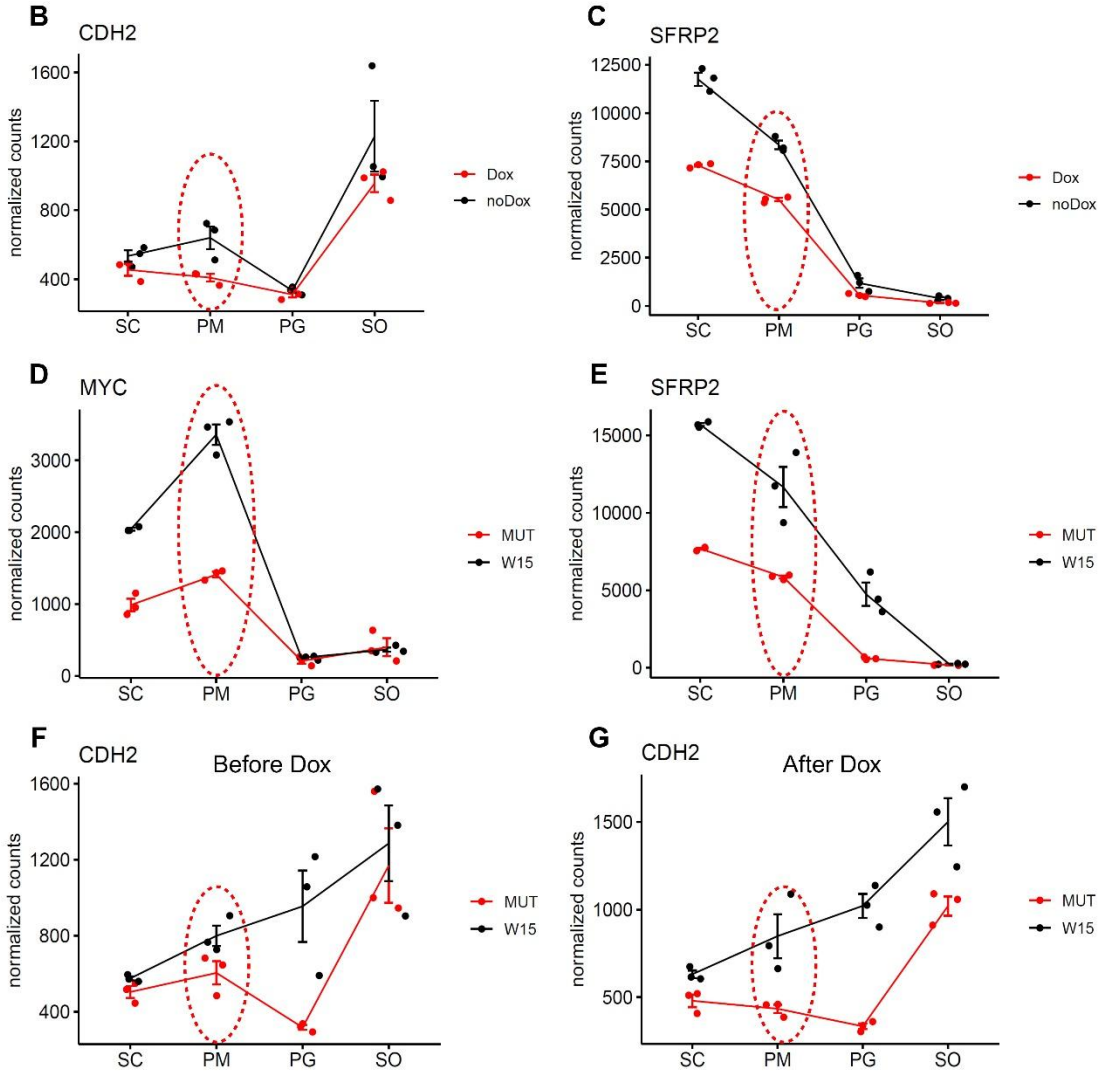
In addition to CDH2 mRNA, MUT-NANOS1 protein impacts the WNT pathway at multiple levels, including binding to WNT modulator SFRP2 via the GAAGAA motif (**Fig. 50A** and **Fig. 50C**). Similarly, MUT-NANOS1 binding reduces SFRP2 mRNA expression compared to the control W15, with the largest difference at the SC and pre-me stages (**Fig. 50E**, time-course analysis). Notably, high levels of SFRP2 in pre-me and SC stages underscore the dual role of WNT signalling, as it supports pluripotency when β-catenin is cytoplasmic (by stabilizing AXIN2) and drives differentiation when β-catenin translocates to the nucleus (Kim et al., 2013).

MYC, another potentially GAAGAA-bound target, is most highly expressed in pre-me cells (**Fig. 50D**), followed by the SC stage, acting as a marker of pluripotency and cell-cycle regulation downstream of the WNT pathway. However, its expression levels drop approximately threefold after MUT-NANOS1 binding compared to the control cell line W15 (**Fig. 50D**).

Together, the loss of MYC, SFRP2, and CDH2 disrupts the balance between WNT-driven proliferation, cell adhesion, and differentiation cues. These findings indicate that MUT-NANOS1's gain-of-function trait, marked by recognition of the SAAGAA (GAAGAA) motif in MYC, SFRP2, and CDH2, is central to its ability to downregulate key WNT pathway components. By binding and repressing these transcripts, MUT-NANOS1 jeopardizes the normal progression from pluripotency toward ME and PGC fates, ultimately compromising germline development.

A

Feature_Type	*Gene Name*	*Motifs*	*Enriched Motif Sequence*	*Counts*	*Sequences %*
CDS	MYC	GAAGAA	SAAGAA	98	16.87
CDS	SFRP2	GAAGAA			
CDS	CDH2	GAAGAA, GAAGAA			



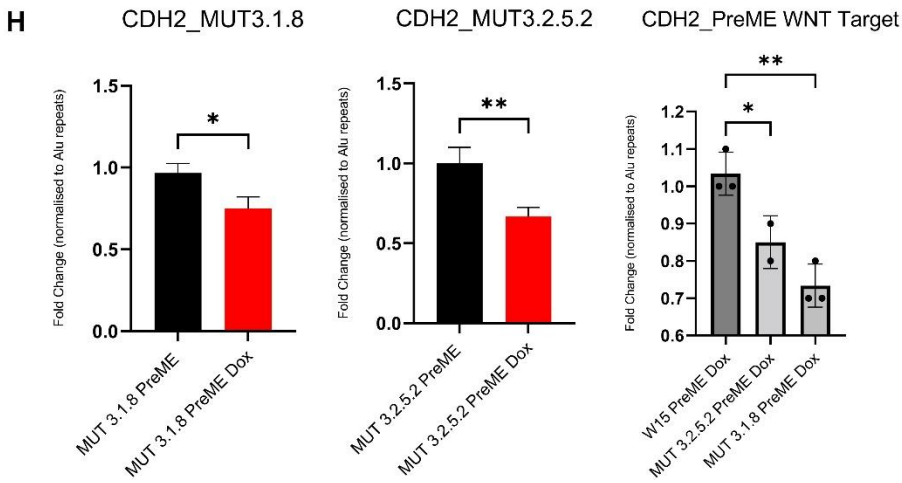


Figure 50. Temporal and motif-dependent modulation of WNT pathway components in response to MUT-NANOS1 overexpression, highlighting the impact on *CDH2*, *SFRP2*, and *MYC* mRNA expression at the pre-me stage and throughout differentiation. A- DREME motif enrichment analysis results, comparing sequences from enriched windows following Skipper analysis of eCLIP-bound data. This analysis identified specific sequence motifs within the CDS of key WNT pathway target and modulator mRNAs including MYC proto-oncogene, bHLH TF (MYC), Secreted Frizzled Related Protein 2 (SFRP2), and Cadherin 2 (CDH2). The motif SAAGAA was enriched in 98 sequences, representing 16.87% of the analyzed sequences. **B-G** ImpulseDE2 time-course analysis of RNA sequencing data, treating developmental stages as time points: stem cell (SC), pre-me (PM), PGC (PG), and soma (SO). In panels **B**- and **C**-, red lines represent dox-treated samples, and black lines represent untreated samples, showing normalized counts of CDH2 and SFRP2, respectively, in dox-treated (dox) versus untreated (no dox) conditions. In panels **D-G**, red lines represent the MUT-NANOS1 cell line 3.2.5.2, and black lines represent the untreated control cell line (W15). Panels **D**- and **E**- display the expression profiles of MYC and SFRP2, respectively, comparing MUT-NANOS1 and control (W15) cell lines following dox induction. Panels **F**- and **G**- show normalized CDH2 expression in the MUT-NANOS1 and W15 cell lines before dox treatment in **F**- and after dox induction **G**. **H**- RT-qPCR analysis of CDH2 expression in both MUT-NANOS1 cell lines (3.1.8 and 3.2.5.2) with and without dox induction, as well as a comparison to control cell line (W15) expression levels (right graph). Expression values (fold changes) are normalized to Alu repeats. Statistical significance was assessed using unpaired one-tailed t-tests (left and middle plots) and one-way ANOVA (right plot with W15 comparison). Asterisks denote significance levels: *$p < 0.05$ and **$p < 0.01$. Standard deviation bars (3 replicates per condition) are shown, and outlier replicates were removed accordingly.

8.1.3 Reduced PGC competence mediated by validated EOMES downregulation following MUT-NANOS1 overexpression in pre-me cells

Finally, it is crucial to evaluate the impact of WNT pathway downregulation on PGC competence, especially since the *in vitro* differentiation protocol aims to generate PGCLCs with the highest possible efficiency. In this context, EOMES, a crucial marker for assessing PGC competence in primed hESCs, was significantly downregulated in both validated MUT-NANOS1 cell lines compared to their own controls (**Fig. 51B**) and to the control cell line W15 (**Fig. 51A**). Time-course analysis showed high expression of EOMES in pre-me cells prior to MUT-NANOS1 expression, (**Fig. 51C**) followed by an approximately 2.3-fold drop (from 1750 to 750 normalized counts) post-dox (**Fig. 51C**), underscoring its stage specificity and in line with the observed WNT suppression.

EOMES is a crucial regulator of PGC competence in pre-me cells, identified through differential expression analyses (Chen et al., 2017; Castillo-Venzor et al., 2023). Pseudo-time trajectory analysis revealed that amnion-like cells (AmLCs) and PGCLCs co-express EOMES, MIXL1, and ZIC transcription factors (Castillo-Venzor et al., 2023). During pre-me and ME stages, early ME enhancers become accessible and acquire H3K27ac, targeting genes involved in the WNT signaling pathway and mesoderm formation, including EOMES, which is essential for SOX17 upregulation during human PGCLC specification (Tang et al., 2022; Kojima et al., 2017). Additionally, motif enrichment analysis indicated that these enhancers are activated by mediators of the FGF and canonical WNT pathways (Kojima et al., 2017). Notably, as a downstream target of the WNT pathway, EOMES is crucial for guiding human PGCLC specification from pre-me precursors and cannot be compensated by TBXT (Alves-Lopes et al., 2023; Kojima et al., 2017).

Overall, premature repression by MUT-NANOS1, via direct binding to target motifs (**Fig. 50A),** likely altered protein interactions, or possible microRNA dysregulation, leads to downregulation of the WNT

signaling pathway in pre-me cells. This disruption lowers EOMES' levels in peri-gastrulation pre-me cells, ultimately reducing PGC competence.

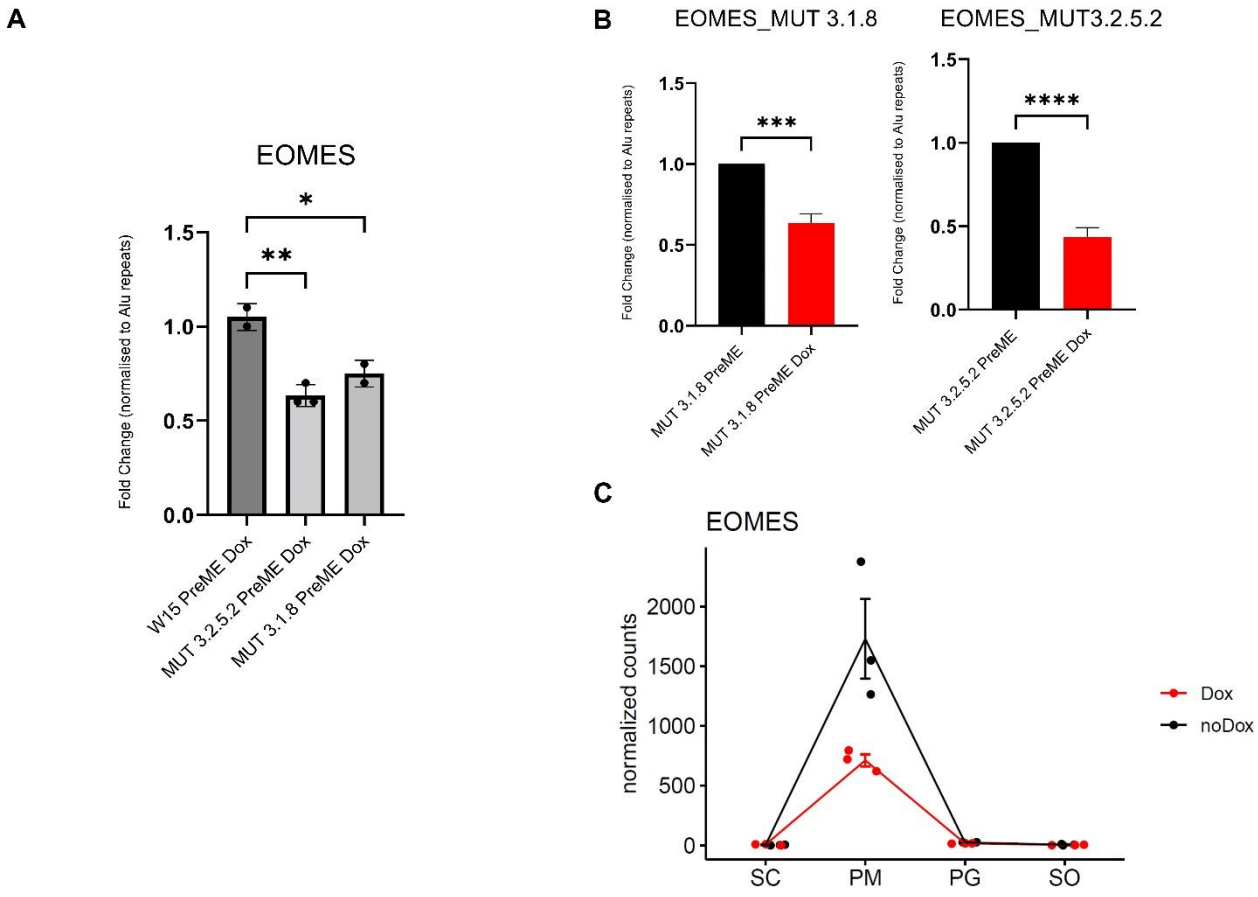


Figure 51. Validation of EOMES expression, a marker of PGC competence in pre-me cells, in response to MUT-NANOS1 overexpression across different developmental stages. A- RT-qPCR analysis of EOMES expression in pre-me cells, comparing control cell line W15 versus two MUT-NANOS1 cell lines (3.1.8 and 3.2.5.2) after dox treatment. Expression values (fold changes) are normalized to Alu repeats. Statistical significance was assessed using one-way ANOVA test. Asterisks denote significance levels: *p < 0.05 and **p < 0.01. (B) RT-qPCR analysis of EOMES expression in two MUT-NANOS1 cell lines (3.1.8 and 3.2.5.2), comparing untreated (black bars, pre-me) and dox-treated (red bars, pre-me dox) conditions within each cell line. Statistical significance was determined using unpaired one-tailed t-tests, with ***p < 0.001 and ****p < 0.0001, indicating significant downregulation of EOMES expression following dox induction in both MUT-NANOS1 cell lines. **C**- ImpulseDE2 time-course analysis of RNA sequencing data, with developmental stages treated as time points: stem cell (SC), Pre-me (PM), PGC (PG), and soma (SO). The plot shows normalized counts of EOMES expression over differentiation stages, with red line indicating dox-treated replicates (dox) and black line representing untreated replicates (no dox). In A and B, standard deviation bars (3 biological replicates per condition) are shown, and outlier replicate/s removed accordingly. Eomesodermin (EOMES).

8.2 Enriched pathways in the altered, MUT-NANOS1 bound RNA interactome of PGCs

In addition to its impact at the pre-me stage, MUT-NANOS1 displays a highly enriched RNA interactome at the PGC stage (**Fig. 42**), with 385 bound and altered targets compared to only 6 for the WT-NANOS1. GO analysis of these targets at the PGC stage of MUT-NANOS1 3.2.5.2 cell line reveals a predominance of

significantly downregulated pathways, based on a minimal p.adjust (q-score) value of 0.005 (**Fig. 52B**), compared to upregulated pathways, which had a minimal p.adjust (q-score) value of 0.06 (**Fig. 52A**).

Among the most significantly downregulated pathways (**Fig. 52B**) are those affecting cell attachment including "focal adhesion" and "cell-substrate junction," followed by "cell adhesion molecule binding" and "cadherin binding." Additionally, pathways related to "protein folding," "response to unfolded protein," and "response to endoplasmic reticulum (ER) stress" were downregulated (**Fig. 52B**) indicating a diminished ability to handle newly synthesized proteins, potentially resulting in cellular stress or apoptosis.

In contrast, the most significantly upregulated pathways (**Fig. 52A**) are primarily associated with structural components, signalling, and morphogenetic processes, indicating enhanced cell differentiation and adaptation. Specifically, increased protein-domain and extracellular matrix (ECM) binding (**Fig. 52A**) point to changes in cell adhesion, migration, and intercellular communication, which are crucial for cell signalling and differentiation cues.

Notably, unlike the pre-me altered RNA interactome (**Fig. 46**), MUT-NANOS1 overexpressing show upregulation of the β-catenin-TCF complex, an essential component of WNT signalling (**Fig. 52A**). This indicates that PGCs exhibit enhanced differentiation or proliferation signals following MUT-NANOS1 expression. TGF-β signaling is also significantly upregulated (albeit less significantly than WNT) (**Fig. 52A**), likely acting downstream to help stabilize WNT signalling. Furthermore, the downregulation of cell adhesion processes in **Fig. 52B** may stem from the elevated WNT and TGF-β signaling, promoting ECM rearrangement and cell reorganization even though the cells are expected maintain their PGC identity.

BMP, WNT, and Nodal (mediated by TGF-β) signalling pathways play distinct and time-sensitive roles in human PGCLCs differentiation. During the 4-day *in vitro* protocol BMP signalling is required continuously for the first two days, whereas WNT operates only within a brief window before the end of Day 1 (Jo et al., 2022). In this study, overexpression of MUT-NANOS1 prematurely suppresses WNT signalling at the pre-me stage prior to BMP4 treatment (**Fig. 46**). Normally, WNT and Nodal pathways are downregulated by days 3 to 4 following BMP4 treatment to enable PGCLC differentiation. However, as shown in **Fig. 52A-B**, MUT-NANOS1 drives upregulation of both WNT and TGF-β (potentially activating Nodal), disrupting the precise timing necessary for PGCLC specification. This disruption may reduce PGC formation, subsequent migration, and gonadal colonization (Heemskerk et al., 2019; Chhabra et al., 2019; Kojima et al., 2017). Proper regulation of Nodal signaling is equally crucial, as premature or prolonged activation can misdirect lineage commitment and hinder PGC development (Jo et al., 2022).

The abrupt changes induced on day 3 of PGCLC differentiation, and the activation of typically dormant pathways, likely demand rapid protein synthesis in response to sudden differentiation cues. This surge may overwhelm the cell's protein quality control mechanisms, as suggested by the diminished ER stress response (**Fig. 52B**). Consequently, misfolded proteins could accumulate, impairing cell function or viability and thereby compromising emerging PGCLCs. Additionally, reduced adhesion may hinder proper cell positioning and communication, disrupting PGC maturation and niche organization in vivo. Overall, the

MUT-NANOS1–altered RNA interactome appears to favor differentiation-related pathways at the expense of protective and adhesion processes, undermining the balance essential for stable PGC identity.

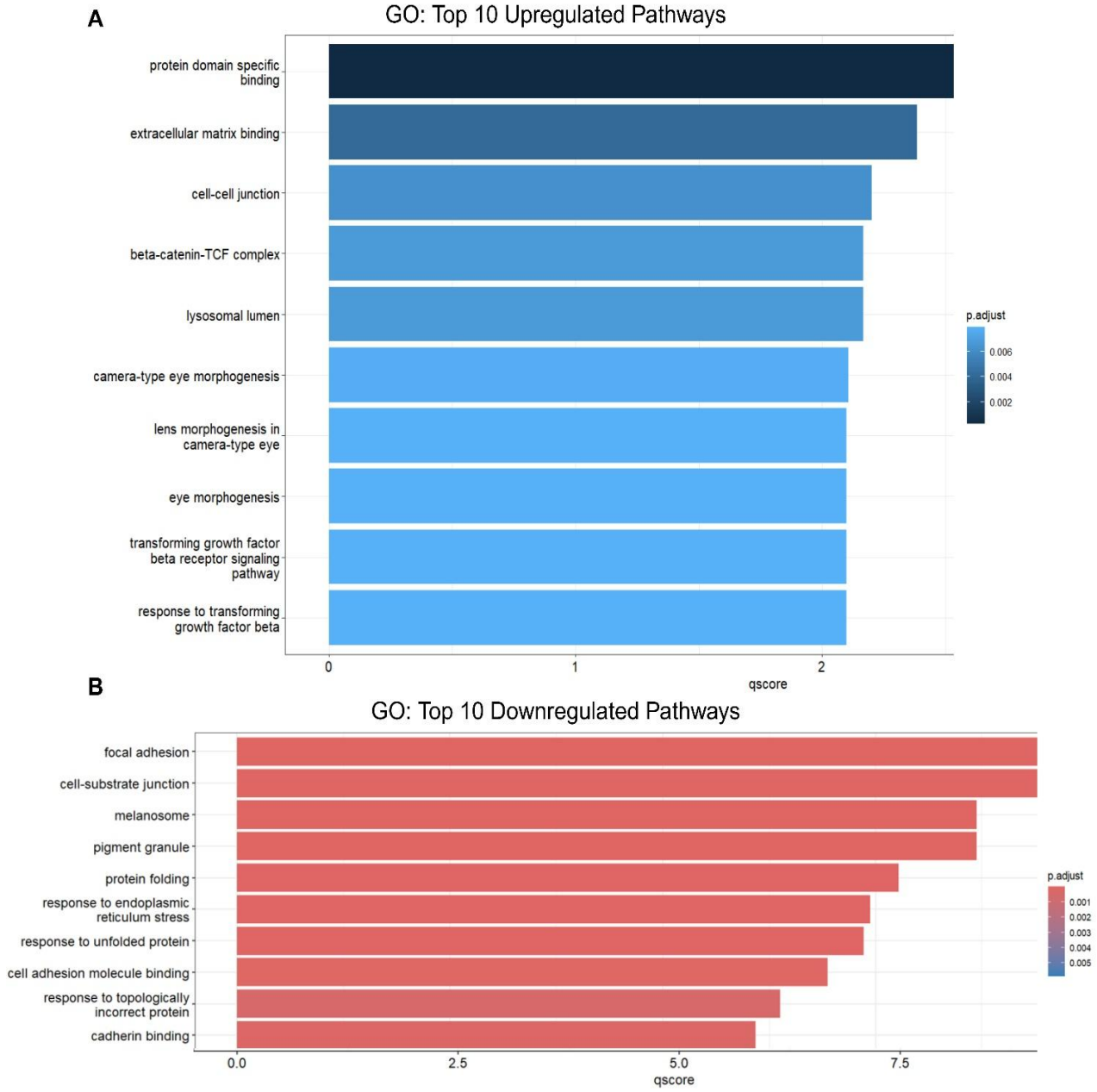


Figure 52. Gene Ontology (GO) analysis depicting the top 10 upregulated and downregulated pathways derived from mRNAs that are both differentially expressed and bound by MUT-NANOS1 at the PGC stage. A- Top 10 upregulated pathways and **B**- top 10 downregulated pathways were identified by integrating eCLIP data, comparing MUT-NANOS1 overexpression before and after dox induction at the pre-me stage with DESeq analysis results from the PGC stage under similar conditions. The top two upregulated pathways (**A**) include "protein domain specific binding" and "extracellular matrix binding," while the top two downregulated pathways (**B**) include "focal adhesion" and "cell-substrate junction." Pathways are ranked by q-score, with adjusted p-values (p. adjust) represented by gradient colour intensity, indicating their significance.

8.2.1 WNT and TGF-β signaling pathway components are differentially expressed in the PGC stage and enriched as MUT-NANOS1-bound transcripts

Building on the GO analysis (**Fig. 52A-B**), a cnetplot (**Fig. 53**) highlights key regulators driving the WNT and TGF-β upregulation, including β-catenin (CTNNB1) and LRP5 mRNAs for WNT, and THBS1 and ADAM9 mRNAs for TGF-β. Notably, BCL9L mRNA uniquely mediates the overlap between both pathways (**Fig. 53**). Two prominently elevated factors are TSKU mRNA, closely linked to WNT, and ZBTB7A mRNA, closely associated with the TGF-β signaling pathway (**Fig. 53**). Importantly, all highlighted mRNAs in **Fig. 53** are bound by MUT-NANOS1as per eCLIP analysis, indicating a significant effect of the MUT-NANOS1 protein on critical differentiation pathways.

ZBTB7A, the most upregulated TGF-β–associated mRNA (**Fig**. **53**), encodes a POK TF with a C-terminal Krüppel-like zinc-finger domain that binds DNA and mediates transcriptional repression by modifying chromatin organization or recruiting additional transcription factors (Gupta et al., 2020). ZBTB7A participates in cell proliferation, differentiation, haematopoiesis, adipogenesis, chondrogenesis, metabolism, and the unfolded protein response. Yu et al. (2020) showed that ZBTB7A binds GACCC sequence motifs and interacts with factors like PAX3, OCT4, and MAFK, remodelling chromatin to activate naïve pluripotency genes while repressing primed genes (e.g., Fgf5, Dnmt3b). This remodelling promotes transitions between pluripotent states.

In the context of PGCs, upregulated ZBTB7A following MUT-NANOS1 overexpression may similarly remodel chromatin, downregulating pluripotency genes and promoting differentiation. BMP signalling influences cell fate decisions, with amnion-like cells exhibiting greater BMP responsiveness than PGCLCs (Castillo-Venzor et al., 2020). The NANOS1 double variant p.[Pro34Thr; Ser78del] might disrupt WT-NANOS1–mediated repression, causing persistent BMP signalling via TGF-β and elevated ZBTB7A. This shift away from a PGC state represses pluripotency markers and activates differentiation pathways, indirectly stimulating WNT signalling, similar to observations in non-small cell lung cancer (NSCLC) (Zhijun and Jingkang, 2017). In NSCLC cells, miR-520e targets ZBTB7A, reducing its expression and modulating WNT signalling, which also remains elevated in MUT-NANOS1–expressing PGCs. Consequently, MUT-NANOS1 binding to ZBTB7A mRNA could sustain BMP and WNT activity, destabilizing the PGC population by restricting pluripotency and advancing differentiation.

TSKU, a small leucine-rich proteoglycan, is the second most upregulated transcript in **Fig. 53** and modulates key signalling pathways involved in cellular communication, proliferation, differentiation, and cell fate determination (Istiaq and Ohta, 2022). In PGCs, TSKU upregulation correlates with WNT pathway activation, evidenced by elevated β-catenin resulting in upregulated differentiation pathway of “eye morphogenesis” (**Fig. 53**), following MUT-NANOS1 overexpression.

Although TSKU can inhibit WNT by competing with WNT2B for FZD4 binding (Ohta et al., 2011; Ito et al., 2021), WNT remains upregulated in the current study, likely due to additional WNT components bound by MUT-NANOS1 and increased TGF-β signaling (**Fig. 53**). Notably, TSKU also plays a dual role in modulating TGF-β signalling in a context-dependent manner, enhancing TGF-β1 expression and activating downstream Smad2/3 phosphorylation in certain tissues (Niimori et al., 2012), which could similarly affect PGCLC differentiation.

Beyond WNT and TGF-β, TSKU influences other developmental pathways (e.g., BMP, Notch, Shh, and IGF) vital for SOX2 expression and proper cell fate decisions (Ahmad et al., 2020). Thus, its overexpression in PGCs may drive inappropriate differentiation, such as "eye morphogenesis," instead of normal PGC progression. By binding to TSKU mRNA, MUT-NANOS1 may insufficiently repress its regulatory role, amplifying differentiation signals while undermining protective and adhesion mechanisms, ultimately reducing PGCLC differentiation efficiency.

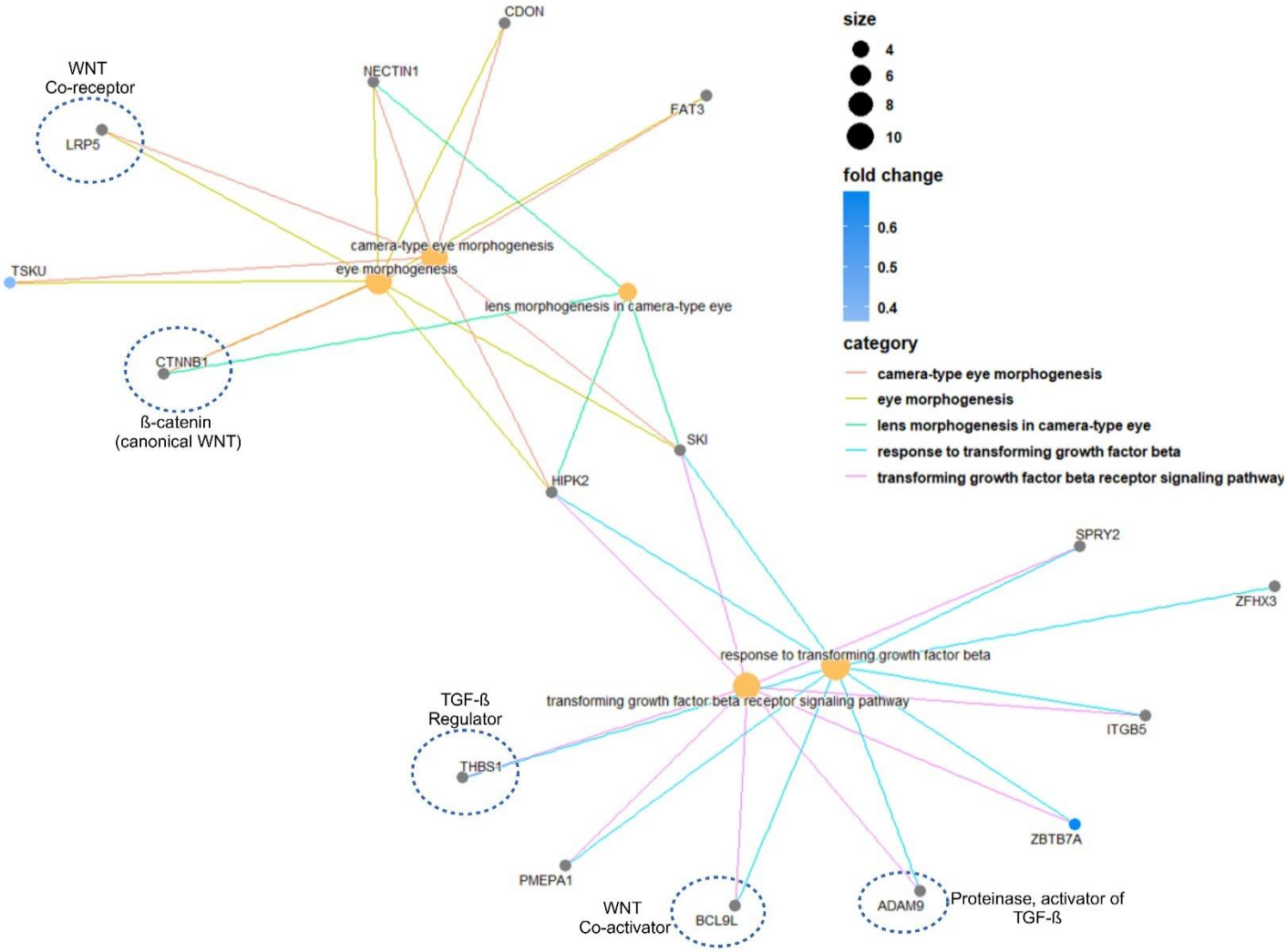


Figure 53. Cnetplot illustrating upregulated WNT and TGF-β signalling pathway mRNAs bound by MUT-NANOS1 in PGCs. This cnetplot illustrates the connections between the top upregulated Gene Ontology (GO) pathways (**Fig. 52A**) and their associated mRNA encoding genes, based on data from the analysis of MUT-NANOS1 overexpression as described in **Fig. 52**. The most upregulated pathways are represented by orange nodes, with node size indicating the number of associated genes. Connecting lines depict gene-pathway relationships, color-coded by pathway category. Gene nodes (gray circles) vary in size according to their significance and are coloured based on fold change (blue gradient), with darker shades representing higher fold changes. Key regulatory genes such as β-catenin (canonical WNT signalling), LRP5 (WNT co-receptor), and THBS1 (TGF-β regulator) are highlighted, underscoring their central roles in linking these pathways.

8.2.2 Time-course and RT-qPCR validation of upregulated WNT pathway transcripts (LRP5, LRP4 and BCL9L) bound by MUT-NANOS1 in PGCs

Initial validation focused on several WNT components from **Fig. 53**, including those downregulated at the pre-me stage (**Fig. 48A-B**). Among these, LRP5, upregulated in **Fig. 53**, was validated *via* RT-qPCR to be significantly increased in both MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 (**Fig. 54A**), corroborating eCLIP and RNA seq. overlapping (**Fig. 52**). Time-course analyses (**Fig. 54B**) show stable LRP5 levels at the pre-me stage, implying a different mechanism for WNT suppression at this stage. In contrast, from the PGCLC sorted population, LRP5 levels increased approximately 1.5-fold following MUT-NANOS1 overexpression (**Fig. 54B**), indicating a stronger WNT-activating role at this stage. LRP5, a critical co-receptor with LRP6, plays an essential role in Wnt/β-catenin signaling and cell fate determination during both germline and somatic development (Liu et al., 2022; Niehrs and Shen, 2010), with mutations in mouse Lrp5 and Lrp6 causing severe gastrulation defects, underscoring their importance in embryogenesis and mesoderm formation (Kelly et al., 2004). In human cells, LRP5 expression increases during differentiation processes, such as osteoblast differentiation, by enhancing canonical WNT signaling (Gong et al., 2001). Notably, LRP5 expression peaks in the somatic population over the time-course analysis (**Fig. 54B**), both before and after MUT-NANOS1 overexpression, suggesting its importance to somatic lineage development and the potential negative impact of MUT-NANOS1–mediated LRP5 upregulation on PGCLC differentiation.

Conversely, LRP4, another member of the low-density lipoprotein receptor family, was downregulated by approximately 2.5-fold at the PGC stage (**Fig. 54D**) and enriched in the RNA interactome bound by MUT-NANOS1 (eCLIP data, not shown). This suppression was significant for both MUT-NANOS1 cell lines (**Fig. 54C**). Unlike LRP5, LRP4 acts as a negative regulator of WNT signaling, opposing the functions of LRP5 and LRP6 (Li et al., 2010). LRP4, structurally similar to LRP5/6 but lacking key WNT co-receptor motifs, serves as a WNT/β-catenin inhibitor (Willnow et al., 2012). By antagonizing LRP5/6 and binding WNT antagonists like DKK and WISE, LRP4 curbs canonical WNT signaling (Johnson et al., 2005; Li et al., 2010; Ahn et al., 2013, 2017). In MUT-NANOS1–expressing PGCs, LRP5 is upregulated (**Fig. 54A-B**), while LRP4 is downregulated (**Fig. 54C-D**), collectively driving persistent canonical WNT activity via β-catenin and TCF/LEF–mediated gene regulation.

Overall, MUT-NANOS1 in PGCs mediates upregulation of the WNT co-receptor LRP5 (**Fig. 54A-B**) and suppresses its inhibitor LRP4 (**Fig. 54C-D**), leading to sustained canonical WNT signaling through β-catenin and TCF/LEF-mediated gene activation.

BCL9L, a WNT co-activator linked to TGF-β signaling (**Fig. 53**), was significantly upregulated in both MUT cell lines 3.1.8 and 3.2.5.2 compared with the W15 control post-dox (**Fig. 54E**). This upregulation, likely involving direct MUT-NANOS1 binding, enhances β-catenin nuclear translocation and transcriptional activity (Ring et al., 2014) and intersects with TGF-β signaling, further modulating cellular processes such as epithelial-mesenchymal transition (EMT). In pancreatic cancer models, increased BCL9L expression facilitates EMT by reducing E-cadherin levels and promoting β-catenin nuclear localization, thereby enhancing cell proliferation and migration (Sannino et al., 2016). Similarly, in the current PGC model, upregulated BCL9L likely amplifies WNT signaling while also influencing TGF-β pathway. This dual effect may disturb the balance required for PGCLC differentiation by favoring differentiation-related processes over pluripotency. Acting as a key intersection between WNT and TGF-β, BCL9L integrates signals from both pathways to guide cellular responses. Moreover, overexpression of MUT-NANOS1 reduces cadherin binding (**Fig. 52B**), further destabilizing WNT signalling and potentially undermining PGCLC differentiation.

Lastly, the roles of β-catenin (CTNNB1) and TCF4 were examined, as both were part of the MUT-NANOS1 RNA interactome and downregulated in pre-me cells (**Fig. 48A-B**; TCF4 downregulation validated via RT-qPCR). While their levels were expected to be upregulated in PGCs, validation via RT-qPCR was challenging due to variability among the three biological replicates. Time-course analyses (**Fig. 54F-G**) revealed that in the MUT 3.2.5.2 cell line prior to MUT-NANOS1 overexpression (no dox, black line), both β-catenin (**Fig. 54F**) and TCF4 (**Fig. 54G**) levels drop markedly from pre-me to the PGC stage, especially TCF4. However, when MUT-NANOS1 was overexpressed (dox, red line), both β-catenin and TCF4 levels increased at the PGC stage, with β-catenin showing a more pronounced rise (**Fig. 54F**), suggesting that direct MUT-NANOS1 binding may reverse the downregulation or not mediate it properly.

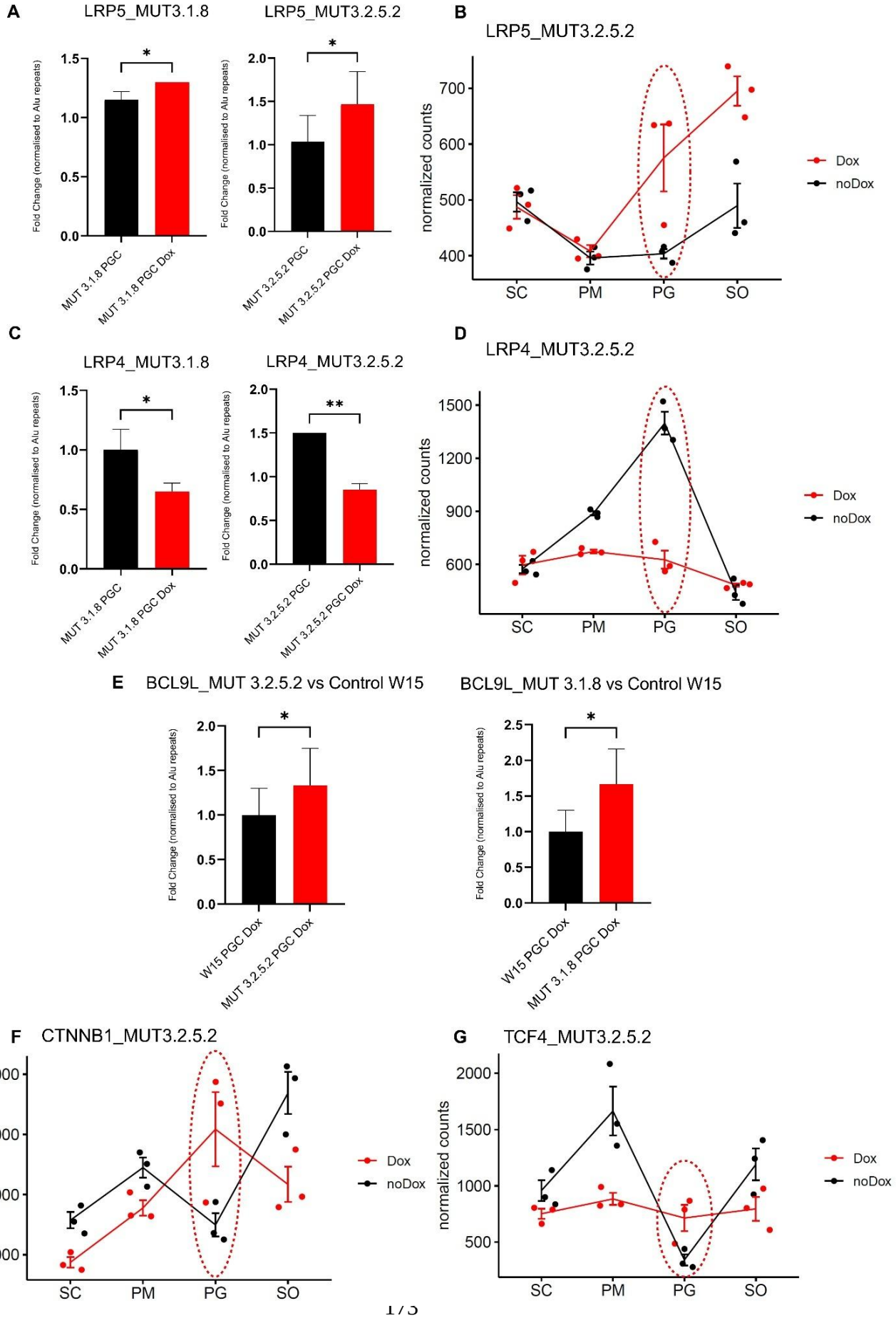

A
LRP5_MUT3.1.8
LRP5_MUT3.2.5.2
Fold Change (normalised to Alu repeats)
MUT 3.1.8 PGC
MUT 3.1.8 PGC Dox
MUT 3.2.5.2 PGC
MUT 3.2.5.2 PGC Dox
B
LRP5_MUT3.2.5.2
normalized counts
SC
PM
PG
SO
Dox
noDox
C
LRP4_MUT3.1.8
LRP4_MUT3.2.5.2
D
LRP4_MUT3.2.5.2
E
BCL9L_MUT 3.2.5.2 vs Control W15
BCL9L_MUT 3.1.8 vs Control W15
W15 PGC Dox
MUT 3.2.5.2 PGC Dox
MUT 3.1.8 PGC Dox
F
CTNNB1_MUT3.2.5.2
G
TCF4_MUT3.2.5.2

Figure 54. Validation and time-course analysis of WNT signaling transcripts LRP5, LRP4, BCL9L, CTNNB1 (β-catenin), and TCF4 in MUT-NANOS1 cell lines. A- Bar plots depicting qPCR fold change of LRP5 in PGC stages for the two MUT-NANOS1 cell lines (3.1.8 and 3.2.5.2). Expression levels of LRP5 are significantly increased after dox addition and MUT-NANOS1 overexpression (*$p < 0.05$; **$p < 0.01$). **B**- Time-course analysis of LRP5 expression across four developmental stages: stem cell (SC), pre-me (PM), PGC (PG), and soma (SO). LRP5 expression peaks at the PGC stage following dox induction (red line), highlighting stage-specific regulation. **C**- Bar plots showing qPCR fold change of LRP4 in PGC stages for the two MUT-NANOS1cell lines. Expression levels of LRP4 are significantly decreased after dox addition and MUT-NANOS1 overexpression (*$p < 0.05$; **$p < 0.01$). **D**- Time-course analysis of LRP4 expression across the same developmental stages. LRP4 expression peaks at the PGC stage prior to dox induction (black line), indicating stage-specific regulation. **E**- Bar plots showing fold change of BCL9L in MUT cell lines compared to the control cell line (W15) after dox treatment and MUT-NANOS1 overexpression in clones. BCL9L is significantly upregulated in both MUT cell lines (*$p < 0.05$). **F**- and **G**- Time-course analysis β-catenin (F) and TCF4 (G), key canonical WNT pathway components, across developmental stages. β-catenin shows upregulation during the PGC stage following dox addition (red line), while TCF4 exhibits less pronounced upregulation after dox induction (red line) compared to the uninduced control (black line), which peaks in PM. For RT-qPCR validation, statistical significance was assessed using an unpaired Student's t-test (one-tailed) with normalization performed relative to Alu repeats as the reference. Time-course analysis was conducted using ImpulseDE, an R package designed for analyzing time-course gene expression data.

8.2.3 RT-qPCR validation of upregulated TGF-β regulators and EMT-associated transcripts (THBS1, CDH1, VCL) bound by MUT-NANOS1 in PGCs

In addition to upregulated WNT components, TGF-β markers were also examined (**Fig. 55C–E**). Notably, THBS1 levels were significantly elevated in both MUT-NANOS1 cell lines compared to the W15 control (**Fig. 55C**). In MUT-NANOS1 3.2.5.2, THBS1 upregulation was statistically significant relative to its dox-uninduced control (**Fig. 55E**), while the 3.1.8 clone showed an upward trend that did not reach significance due to replicate variability (**Fig. 55D**).

THBS1, a key member of the thrombospondin family, is a multifunctional matricellular protein that increases during tissue injury and repair (Adams and Lawler, 2004). It serves as a major activator of the TGF-β1 pathway, promoting EMT and tumour invasion and migration (Massagué, 2012). However, its ability to activate TGF-β1 varies across tumour types, as seen in glioblastoma cells where THBS1 does not enhance TGF-β1 signaling (Daubon et al., 2019). Additionally, TGF-β1 signaling upregulates THBS1 via SMAD3, establishing a positive feedback loop that further amplifies the pathway (Daubon et al., 2019). In melanoma cells, THBS1 promotes invasion by activating TGF-β1, highlighting its role in EMT and tumour progression (Jayachandran et al., 2014).

EMT is characterized by the downregulation of E-cadherin (CDH1) and the upregulation of N-cadherin (CDH2), facilitating cellular motility (Polyak & Weinberg, 2009). In the PGC model, GO analysis upon MUT-NANOS1 overexpression revealed significant downregulation of "cadherin binding," "focal adhesion," and "cell adhesion molecule binding" (**Fig. 52B**). Accordingly, the levels of CDH1 and vinculin (VCL), components of adherens junctions (AJs), were validated via RT-qPCR and were found to be significantly downregulated in both MUT-NANOS1 cell lines compared to their respective uninduced controls (**Fig. 55A-B**). In the context of PGCs, the aberrant upregulation of THBS1 (**Fig. 55C-E**) and downregulation of CDH1 (**Fig. 55A**) and VCL (**Fig. 55B**) mirror the EMT-

like processes observed in cancer cells. By binding to the mRNAs of THBS1, CDH1, CDH2, and VCL, MUT-NANOS1 disrupts their normal expression (**Fig. 45** and **Fig. 48**), impairing the interaction between β-catenin and cadherins. Normally, CDH1 links β-catenin to the cell surface, stabilizing AJs and maintaining cell adhesion. The loss of CDH1 as well as VCL disrupts these junctions, allowing β-catenin to engage in canonical TCF-mediated WNT signaling and promote EMT (Huber et al., 2001; Yap et al., 1997; Wal & van Amerongen, 2020; Peng et al., 2010). Additionally, downregulation of VCL, which anchors AJs to the cytoskeleton, further weakens cell adhesion. These disruptions enhance WNT signaling and TGF-β-mediated EMT, altering cell fate decisions and hindering PGCLC differentiation.

Overall, overexpression of MUT-NANOS1 in PGCs results in coordinated dysregulation of THBS1, CDH1, and VCL transcripts, a pattern characteristic of EMT-like processes in cancer cells. This abnormal regulation indicates that MUT-NANOS1 may influence cell behavior through similar pathways to those driving tumour invasion, ultimately disrupting PGC identity and redirecting cells toward a somatic or mesoderm-like lineage.

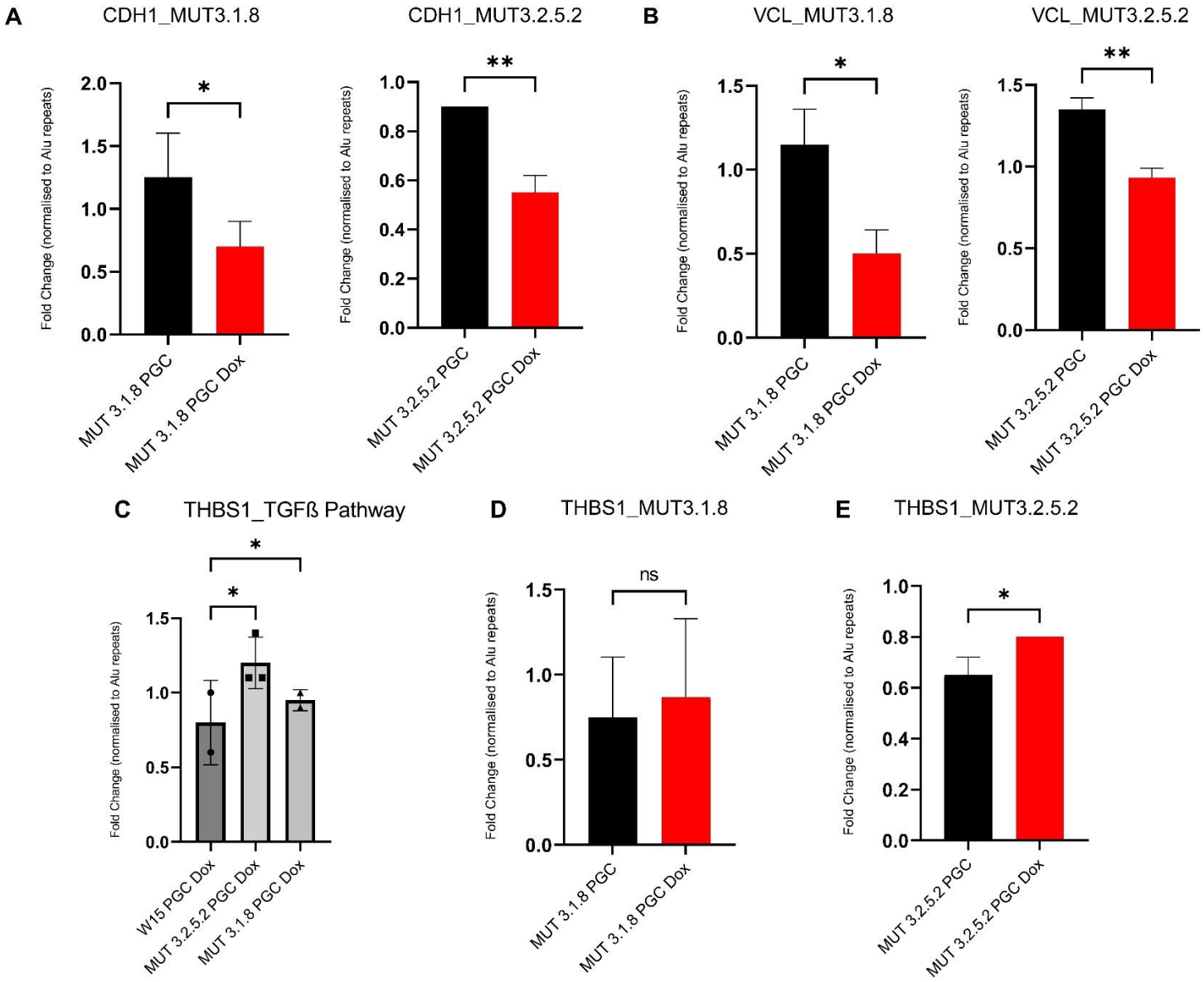


Figure 55. Validation of Thrombospondin 1 (THBS1), E-Cadherin (CDH1), and Vinculin (VCL) expression in MUT-NANOS1 cell lines at the PGC Stage. RT-qPCR results are presented to validate the expression of THBS1 (a regulator of the TGF-β signaling pathway) and epithelial-to-mesenchymal transition (EMT)-associated markers E-Cadherin (CDH1) and VCL in two MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2. **A**- CDH1 expression in both MUT cell lines shows a significant reduction following dox-induced overexpression of MUT-NANOS1 (*p < 0.05; **p < 0.01). **B**- VCL expression is significantly reduced in both MUT-NANOS1 3.1.8 and 3.2.5.2 cell lines after dox induction of MUT-NANOS1 (*p < 0.05; **p < 0.01). **C**- THBS1 expression is significantly upregulated in both MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 compared to the dox-treated W15 control PGCLCs (W15 PGC dox) (*p < 0.05; one-way ANOVA). **D**- THBS1 expression in MUT 3.1.8 PGCLCs is upregulated compared to its uninduced control; however, this increase is not statistically significant. **E**- THBS1 expression in MUT 3.2.5.2 cell line is significantly upregulated after dox treatment compared to its uninduced control (*p < 0.05). All fold changes were normalized to Alu repeats. Statistical analyses were performed using an unpaired Student's t-test (one-tailed) for comparisons between two groups and one-way ANOVA for panel (C).

9. Upon overexpression, MUT-NANOS1 cell lines exhibit reduced levels of PGC markers

Following the validation of bound mRNAs involved in the aberrant upregulation of the WNT and TGF-β signaling pathways in PGCLCs (**Fig. 54A-G**; **Fig. 55A-E**), the expression levels of classical PGC markers were examined in both MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2. This analysis aimed to determine whether the overexpression of MUT-NANOS1 and the consequent aberrant signaling had a significant effect on the PGC-like characteristics expected of the differentiated PGCLCs collected four days post-differentiation.

The specification and maintenance of PGCs rely on a tightly regulated network of transcription factors and markers, which is disrupted by the overexpression of MUT-NANOS1, leading to the downregulation of several key germ cell markers (**Fig. 56**). Among these, SOX17 plays a pivotal role as the primary regulator of human PGCLC specification by activating genes essential for germ cell identity (Irie et al., 2015). In MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2, SOX17 expression is significantly reduced (**Fig. 56**). While the reduction in the MUT-NANOS1 3.1.8 cell line is statistically significant, it shows slightly lower significance due to high variability among three biological replicates. In contrast, the MUT-NANOS1 3.2.5.2 cell line exhibits a more pronounced and statistically significant downregulation (**Fig. 56**). Similarly PRDM1, a critical factor that represses somatic and endodermal genes during germ cell specification (Irie et al., 2015), is also downregulated (**Fig. 56**). Functionally, PRDM1 acts downstream of SOX17 to mediate this repression, facilitating germ cell specification (Irie et al., 2015). Unlike SOX17 validation, while PRDM1's downregulation is statistically significant in the MUT-NANOS1 3.2.5.2 cell line, the variability among the tested replicates in MUT-NANOS1 3.1.8 differentiated cell line limits conclusive results, emphasizing the need for additional biological replicates to confirm the findings (**Fig. 56**).

PRDM14, while critical for mouse PGC specification, exhibits delayed and significantly reduced expression in human PGCLCs, with lower levels detected in TCam-2 seminoma cells compared to hESCs (Grabole et al., 2013; Ma et al., 2011). In MUT-NANOS1 cell lines, PRDM14 is significantly downregulated in MUT-NANOS1 3.2.5.2 cell line, while MUT-NANOS1 3.1.8 PGCs show reduced expression with high replicate variability, resulting in a lack of statistical significance, similar to PRDM1 (**Fig. 56**). Notably, PRDM14 marker is essential for maintaining pluripotency and facilitating the transition from hESCs to PGCs, with its precise regulation being pivotal for PGC differentiation (Sybirna et al., 2020). It was shown that its depletion disrupts PGC specification and causes transcriptome changes similar to those observed with the loss of TFAP2C or PRDM1 (Sybirna et al., 2020).

TFAP2C and NANOS3, crucial markers for human PGC identity, are consistently and significantly downregulated in both MUT-NANOS1 cell lines (**Fig. 56**) unlike PRDM1 and PRDM14. Notably, TFAP2C and NANOS3 play vital roles in maintaining PGC identity and ensuring proper

migration and gonadal colonization (Kojima et al., 2017; Sasaki et al., 2015; Li et al., 2017). Their coordinated downregulation highlights the disruption of the PGC specification pathway caused by MUT-NANOS1 overexpression. TFAP2C works alongside SOX17 and PRDM1 to regulate key germ cell identity genes, including NANOS3, OCT4 (POU5F1), and NANOG (Tang et al., 2022).

Unsurprisingly, both OCT4 and NANOG mRNA levels were also significantly downregulated in both MUT-NANOS1 sorted PGCLCs (**Fig. 56**), likely due to reduced levels of upstream regulators SOX17, PRDM1, and TFAP2C, the latter also promoting chromatin reorganization through H3K9 demethylases (Eguizabal et al., 2016). Furthermore, in collaboration with PRDM1, TFAP2C also regulates chromatin remodeling components such as SMARCA2 and ARID1B (Alver et al., 2017) and represses somatic genes linked to embryonic development and differentiation (Tang et al., 2022).

On the other hand, the downregulation of NANOS3, directly regulated by the cooperative actions of SOX17, TFAP2C, and PRDM1 (Tang et al., 2022), underscores the failure of upstream TFs to maintain PGC-specific gene expression. As a key downstream target, NANOS3 supports PGC survival and maintenance by preventing apoptosis and reinforcing the germ cell program (Tang et al., 2022).

Collectively, SOX17, TFAP2C, and PRDM1 integrate signals to regulate PGC-specific genes, repress somatic genes, and establish a chromatin landscape conducive to human PGC fate (Tang et al., 2022). Furthermore, SOX17 and TFAP2C drive human PGC fate by shaping a chromatin landscape favorable for PGC identity, initiating the expression of PRDM1 and NANOS3, and upregulating early markers like WNT2 and NODAL (Tang et al., 2022; Castillo-Venzor et al., 2023).

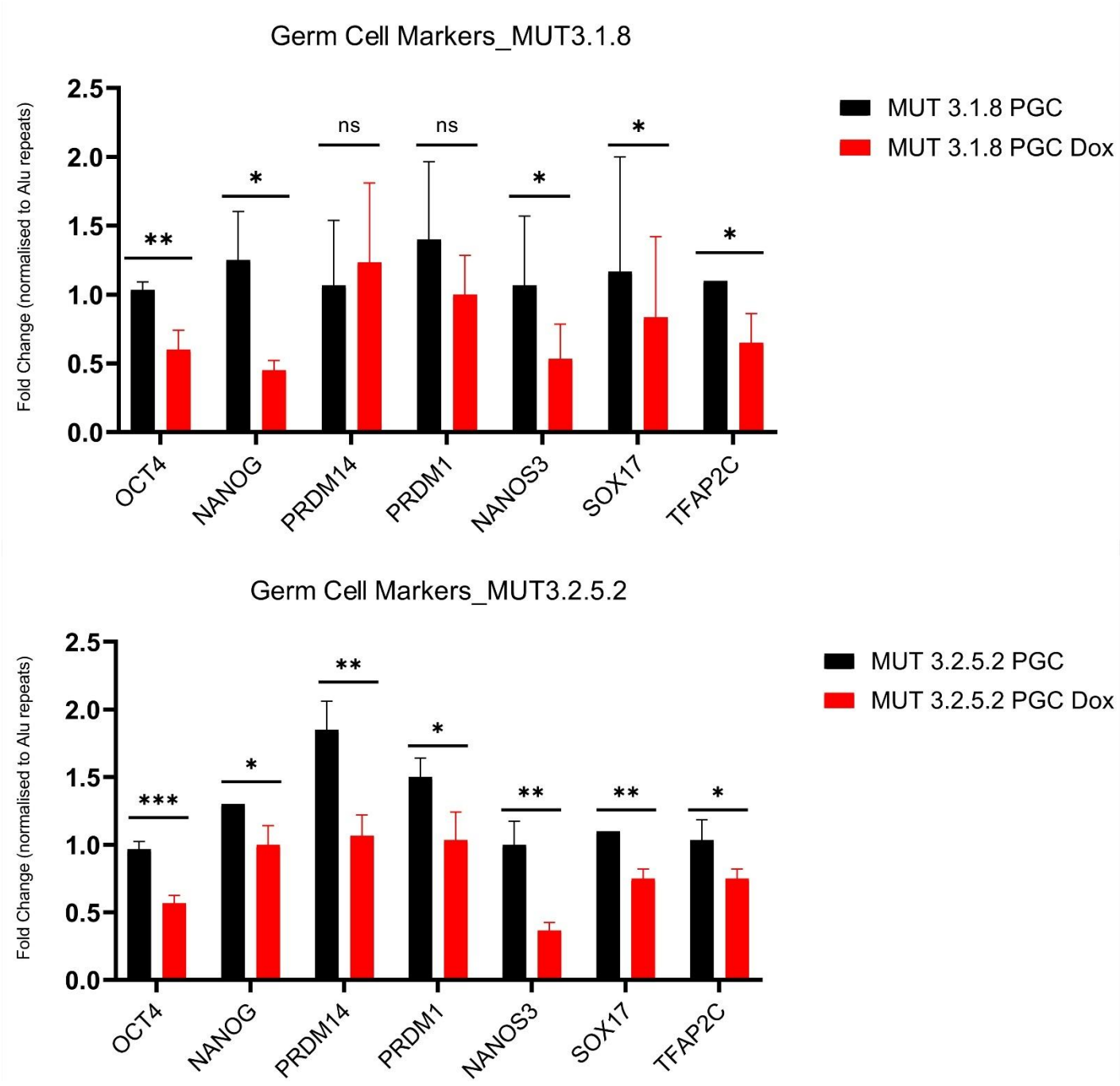


Figure 56. Validation of PGC markers levels in MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 before and after dox induced overexpression of MUT-NANOS1 at the PGC stage. RT-qPCR analysis was performed to assess the expression of PGC markers before (black bars) and after (red bars) dox induction to overexpress the MUT-NANOS1 protein. Expression levels were normalized to Alu repeats, and statistical significance was determined using multiple unpaired one-tailed Student's t-tests. Top graph (MUT 3.1.8): Germ cell markers including OCT4, NANOG, NANOS3, SOX17, and TFAP2C exhibited significant downregulation following dox addition (*$p < 0.05$; **$p < 0.01$), suggesting a reduction in PGC identity. Bottom graph (MUT 3.2.5.2): All assessed markers such as OCT4, NANOG, PRDM14, PRDM1, NANOS3, SOX17, and TFAP2C were all significantly downregulated (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$) upon dox treatment, indicating a more pronounced repression of PGC markers compared to MUT-NANOS1 3.1.8 cell line.

10. Binding of MUT-NANOS1 drives downregulation of pluripotency factors OCT4 and NANOG

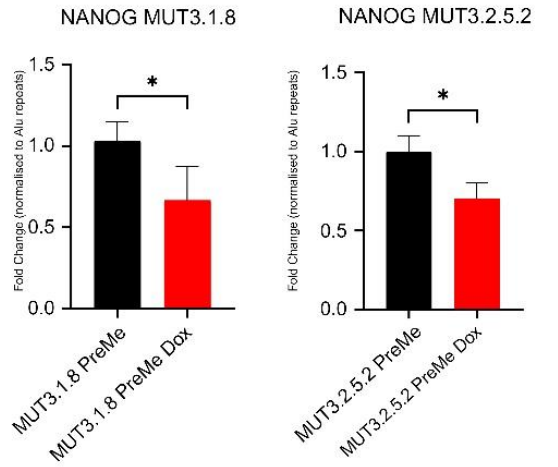
Pluripotency factors OCT4 and NANOG were further investigated because their significant downregulation at the PGC stage (**Fig. 56**) was not only mediated by the reduction of upstream TFs but also due to potential direct binding by MUT-NANOS1 (**Fig. 57C**). eCLIP analysis showed that MUT-NANOS1, unlike its WT counterpart, binds to the 5′UTR of NANOG mRNA and the 3′UTR of OCT4 mRNA (**Fig. 57C**). These bindings could potentially play a key role in post-transcriptional regulation, impacting mRNA stability, localization, and translation, which correlates with OCT4 and NANOG altered mRNA levels observed in **Fig. 56**, **Fig. 57A-B** and **Fig. 57D-E**. In addition to the PGC stage, the mRNA levels of OCT4 and NANOG were quantified at the pre-me stage, where they were significantly downregulated in both MUT-NANOS1 cell lines following MUT-NANOS1 overexpression (**Fig. 57A** for NANOG and **Fig. 57B** for OCT4). Therefore, these crucial pluripotency factors for PGC specification are likely directly affected by MUT-NANOS1 binding, which induces aberrant downregulation of both NANOG and OCT4 levels initiated at the pre-me stage and continuing into the PGC stage.

Notably, OCT4 and NANOG are essential TFs that regulate pluripotency during early embryogenesis and in iPSCs (Irie et al., 2014). Their coordinated expression is critical for PGC competence, guiding PGC specification and development into PGCLCs. In pig pre-gastrulation embryos, OCT4 is uniformly expressed across the epiblast at E13, with some posterior cells co-expressing OCT4 and NANOG, likely marking PGC precursors. By E15, after the primitive streak forms, NANOG becomes restricted to OCT4-positive PGCs at the posterior pole (Irie et al., 2014). Similarly, human migratory PGCs and gonocytes express OCT4 and NANOG, akin to mouse PGCs. *In vitro*, pluripotent stem cells forming alkaline phosphatase-positive colonies also express OCT4, SOX2, and NANOG (Irie et al., 2014).

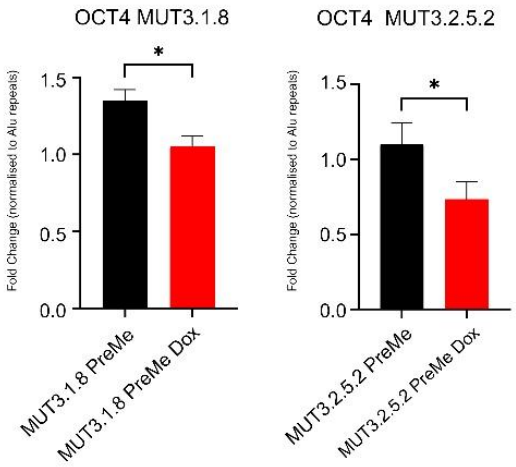
These studies, reflecting the presence and role of OCT4 and NANOG from stem cells to PGCs, are consistent with the MUT 3.2.5.2 cell line time-course analysis for NANOG (**Fig. 57D**) and OCT4 (**Fig. 57E**). In this cell line, NANOG and OCT4 levels steadily increase from the stem cell stage to pre-me, peak at the PGC stage, and then decline in the somatic replicates (**Fig. 57D-E**). However, following MUT-NANOS1 overexpression (red line, dox), both markers decrease at all differentiation stages, most notably at the PGC stage, while remaining unchanged in the somatic population where they are not required (**Fig. 57D-E**). This pattern suggests that MUT-NANOS1 likely represses these mRNAs directly, with inhibition beginning at the stem cell stage, intensifying in pre-me, and peaking at the PGC stage, where these factors are most critical. Normally, during this differentiation window, SOX17, TFAP2C, and PRDM1 collaborate with OCT4 and NANOG to regulate germline enhancers that are already primed in hESCs and pre-me cells and activated in PGCLCs (Tang et al., 2022). Notably, enhancer sequences for OCT4 and NANOG are regulated by upstream SOX17 and TFAP2C, both of which are downregulated following MUT-NANOS1 overexpression (**Fig. 56**).

Based on the findings in **Fig. 56** and **Fig. 57A-E**, overexpression of MUT-NANOS1 during the pre-me and PGC stages leads to downregulation of OCT4 and NANOG mRNAs, likely through binding the NANOG 5′UTR and OCT4 3′UTR, thereby disrupting mRNA stability and translation. Since OCT4 and NANOG are essential for maintaining pluripotency and driving PGC-specific gene expression, their repression impairs the regulatory network crucial for PGC competence and differentiation. Consequently, reduced cooperation among OCT4, NANOG, SOX17, and TFAP2C hinders germline enhancer activation and chromatin remodelling, ultimately disrupting human PGC fate and germ cell functionality.

A



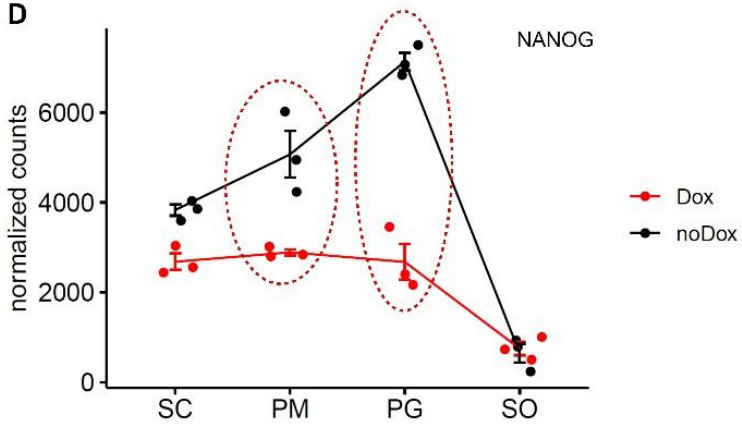
B



C

Chr	Start	End	Strand	Feature type	Top gene name
12	7789543	7789614	+	5'UTR	NANOG
6	31164424	31164512	-	3'UTR	OCT4

D



E

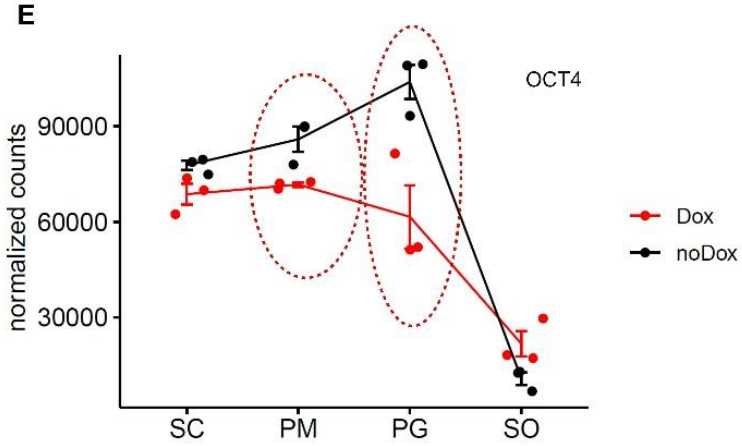


Figure 57. Temporal expression analysis and validation of NANOG and OCT4 mRNA levels in MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2. A- RT-qPCR analysis of NANOG in MUT 3.1.8 and 3.2.5.2 in pre-me. NANOG expression levels are significantly reduced in the pre-me stage after dox-induced overexpression of MUT-NANOS1 (*$p < 0.05$; unpaired Student's t-test, one-tailed). **B**- RT-qPCR analysis of OCT4 in MUT 3.1.8 and MUT 3.2.5.2 in pre-me. OCT4 expression is significantly decreased in both MUT cell lines following dox addition (*$p < 0.05$). **C**- Genomic binding sites of MUT-NANOS1. The table highlights the genomic binding sites of the MUT-NANOS1 protein at the pre-me stage, showing enrichment in the 5'UTR of NANOG mRNA and the 3'UTR of OCT4 mRNA. **D**- Temporal expression of NANOG across four developmental stages: stem cell (SC), pre-me (PM), PGC (PG), and soma (SO). NANOG expression shows a marked reduction during the PM and PG stages after dox treatment (red line) compared to untreated controls (black line). **E**- Temporal expression of OCT4 exhibits significant repression in the PM and PG stages following dox induction (red line) compared to untreated controls (black line). All expression levels were normalized to Alu repeats. Statistical significance was assessed using unpaired one-tailed Student's t-tests.

11. Temporal expression patterns of OCT4 and NANOG in WT-NANOS1 and MUT-NANOS1 cell lines before and after dox induction

Further analysis revealed that WT-NANOS1 does not bind pluripotency markers OCT4 and NANOG (eCLIP data, WT-NANOS1 RNA interactome). Nonetheless, DESeq analysis showed that NANOG was significantly downregulated in the WT 1.2.5.1 cell line following WT-NANOS1 overexpression. This downregulation is also reflected in the time-course analysis across different stages of differentiation for the WT cell line before (black line, no dox) and after (red line, dox) dox addition. Specifically, NANOG levels dropped by approximately two-fold at the PGC stage, while the downregulation at the pre-me stage was less pronounced for this cell line (**Fig. 58E**). In contrast, OCT4 did not exhibit significant downregulation following WT-NANOS1 overexpression in either pre-me or PGC stages based on DESeq analysis. Consistently, in the time-course analysis, OCT4 expression remained largely unchanged across both pre-me and PGC stages (**Fig. 58F**).

Notably, prior to NANOS1 overexpression, the MUT cell line 3.2.5.2 displayed higher levels of pluripotency markers NANOG (**Fig. 58A**) and OCT4 (**Fig. 58B**) than the WT cell line 1.2.5.1, particularly at the PGC stage and to a lesser extent at the pre-me stage. These findings are also supported by immunofluorescence at the stem cell stage (**Fig. 18**), where the WT-NANOS1 cell line showed lower OCT4, NANOG, and SOX2 protein levels compared to the MUT-NANOS1 cell line in the uninduced state.

After dox induction, NANOG levels became similar between MUT and WT cell lines in both the pre-me and PGC stages (**Fig. 58C**), reflecting a relative downregulation compared to their own uninduced controls (see **Fig. 57D** for MUT cell line and **Fig. 58E** for WT cell line). A similar pattern was observed for OCT4, although its expression was slightly lower in the MUT-NANOS1 cell line compared to the WT counterpart at the PGC stage, while at the pre-me stage, the levels were comparable, though the WT-NANOS1 replicates exhibited high standard deviation (**Fig. 58D**).

Given the higher pluripotency profile of cell line MUT 3.2.5.2 compared to WT 1.2.5.1 before dox induction (**Fig. 18** and **Fig. 58A-B**), the significant downregulation of NANOG and OCT4 in MUT 3.2.5.2 can be primarily attributed to MUT-NANOS1 protein binding and inducing downregulation,

rather than to cellular stress resulting from dox addition and protein overexpression via PiggyBac transposition.

In contrast, the lower pluripotency profile of the WT 1.2.5.1 cell line following genetic editing did not improve upon overexpression of WT-NANOS1, particularly regarding OCT4 and NANOG levels (**Fig. 51E-F**). This suggests that endogenous NANOS1 may not play a major role at this early stage of PGC specification from pre-me precursors to PGCs prior to migration, and therefore its overexpression does not elicit a significant restorative response. Moreover, the addition of dox and the overexpression of WT-NANOS1 may have negatively impacted the pluripotency of the cell line, especially concerning NANOG levels (**Fig. 58E**). NANOG has been shown to play a crucial role in the proliferation and survival of migrating PGCs, acting as a safeguard of the PGC-specific molecular network in mouse studies (Yamaguchi et al., 2009). The effect of NANOG downregulation on PGC numbers became apparent even when OCT4 was adequately expressed (Yamaguchi et al., 2009), paralleling the findings in **Fig. 58E-F**, where OCT4 levels did not differ significantly post-dox addition, but NANOG levels were more notably affected, particularly at the PGC stage. The pre-me and PGCs in the WT cell line, which were already differentiating at a lower pluripotency level, likely did not benefit from the overexpression of WT-NANOS1 at either the pre-me or PGC stage. Instead, the cells in each respective stage may have undergone a stress response, further diminishing their pluripotency profile, potentially leading to apoptosis, given the more significant effect on NANOG levels (**Fig. 58E**). Alternatively, the duration of overexpression (6 h at pre-me and 24 h at PGC stage) might have been insufficient to mediate any positive effects expected from WT-NANOS1 overexpression. However, given the lower initial pluripotency profile of the WT cell line, it is unlikely that a longer overexpression period would have induced the desired ameliorative effects.

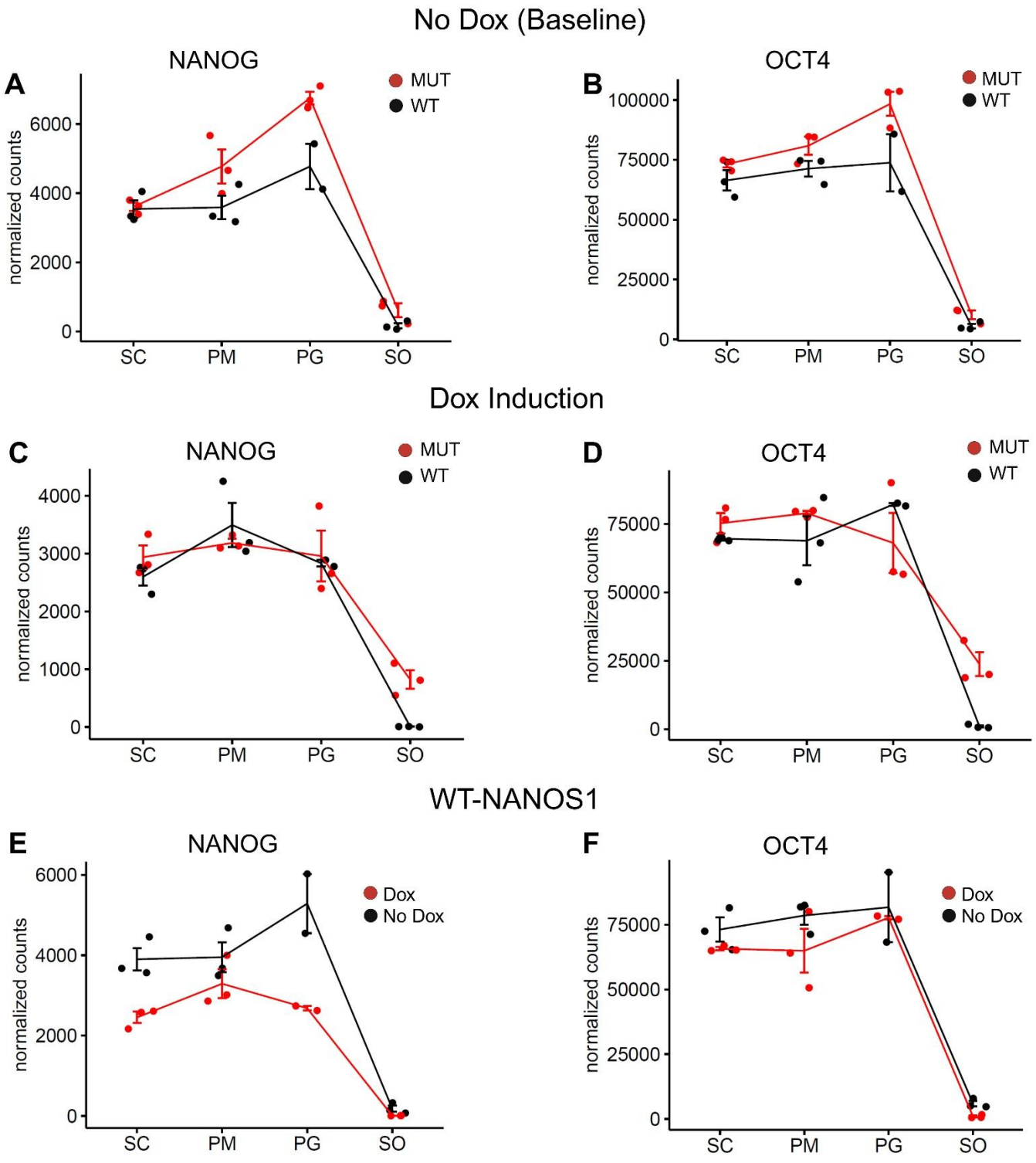


Figure 58. Time-course analysis of NANOG and OCT4 expression in WT 1.2.5.1 and MUT 3.2.5.2 cell lines. **A**- Before dox induction, NANOG levels are higher in the MUT cell line (red line) compared to the WT counterpart (black line) during the PM and PG stages. **B**- Before dox induction, a similar trend is observed for OCT4, with higher levels in the mut cell line during the PM and PG stages. **C**- After dox induction, NANOG levels significantly decrease in both cell lines with the MUT-NANOS1 cell line showing a slightly greater reduction in the PM stage and comparable levels in the PG stage compared to WT-NANOS1 cell line. **D**- After dox induction, OCT4 levels exhibit a minor decrease in the WT cell line and a slightly greater downregulation in the MUT cell line, especially in the PG stage, suggesting more stable OCT4 expression in WT-NANOS1

PGCLCs. **E**- Temporal expression analysis of NANOG mRNA levels in WT 1.2.5.1: Dox treatment (red line) exacerbates the decline in NANOG expression during the PM and PG stages compared to uninduced controls (black line). **F**- Temporal expression analysis of OCT4 mRNA levels, which are less affected following dox induction in WT cell line, maintaining relatively stable levels across all stages aside from a slight decrease in the PM stage. Stem Cell (SC), pre-me (PM), PGC (PG), soma (SO).

12. Significant downregulation of NANOS3 mRNA levels in MUT-NANOS1 cell lines leads to reduced PGC numbers *in vitro* following dox induction

In light of the validated downregulation of classical PGC and pluripotency markers in both MUT-NANOS1 cell lines (**Fig. 56**), it was essential to assess the impact on the differentiation efficiency of PGCs *in vitro*. NANOS3 mRNA, which was significantly downregulated (**Fig. 56**), also served as a marker for analyzing and sorting differentiated cells into PGCLCs and somatic cells.

Upon conducting ten separate differentiation trials, a significant reduction in PGC differentiation efficiency was observed exclusively in the MUT 3.2.5.2 cell line following MUT-NANOS1 overexpression (**Fig. 59**, top left and bottom graphs). Two statistical analyses were performed: The first compared the percentages of NANOS3-tdTomato fluorescence (of uninduced control replicates) and double NANOS3-tdTomato and NANOS1-GFP positive cells in dox-induced MUT 3.2.5.2 replicates (**Fig. 59**, top left graph).

The second analysis employed a two-way ANOVA incorporating both WT and MUT cell lines differentiation before and after dox induction, with cell line and protein overexpression (via dox induction) as co-factors taken into account. This analysis also confirmed that the downregulation in differentiation efficiency was significant only for the MUT-NANOS1 cell line (**Fig. 59**, bottom graph).

The WT 1.2.5.1 cell line exhibited a slight decrease in differentiation efficiency (**Fig. 59**, top right and bottom graphs), which could be attributed to the stress inducing overexpression system and the cell line's baseline lower pluripotency. However, this reduction was not statistically significant, possibly because the downregulation was mediated solely by the cell line factor rather than by the overexpressed WT-NANOS1 protein.

These findings demonstrate that overexpression of MUT-NANOS1, but not the WT counterpart, leads to a significant reduction in PGC differentiation efficiency, even when accounting for the lowered pluripotency of the WT cell line (**Fig. 59**, top left and bottom graphs). The aberrant binding and temporally misaligned repressive signaling of the MUT-NANOS1 protein result in the downregulation of PGC and pluripotency markers (**Fig. 56**), ultimately reducing the differentiation efficiency of human PGCs *in vitro* (**Fig. 59**).

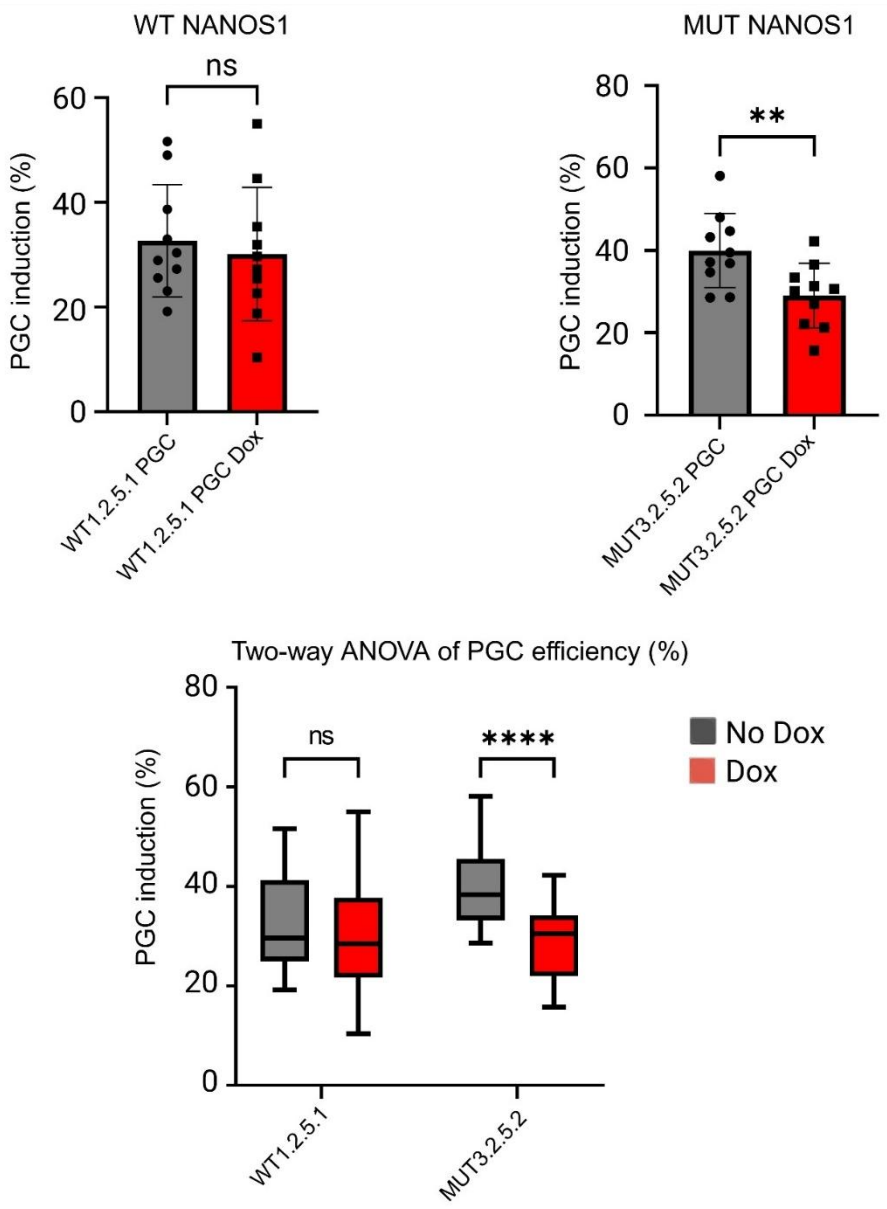


Figure 59. ***In vitro*** **PGC induction efficiency in WT 1.2.5.1 and MUT 3.2.5.2 NANOS1 cell lines.** The induction efficiency of PGCs was quantified across 10 separate differentiations using flow cytometry. This measurement was based on the percentage of td-Tomato fluorescence prior to dox induction. Fluorescence assessments include the combined signal from td-Tomato and double-positive NANOS3-tdTomato and NANOS1-sfGFP populations in dox-induced replicates. Statistical analyses were conducted using unpaired two-tailed Student's t-tests for two-group comparisons and two-way ANOVA considering both cell line and dox induction as co-factors. (Top left graph) The MUT 3.2.5.2 cell line exhibited a significant decrease in PGC induction efficiency following MUT-NANOS1 overexpression (red bar) (**$p < 0.01$), indicating a negative impact on PGC specification. (Top right graph) The WT 1.2.5.1 cell line showed no significant change in PGC induction efficiency after WT-NANOS1 overexpression (red bar), reinforcing that WT-NANOS1 does not notably influence PGC formation at this stage. (Bottom graph) Two-way ANOVA analysis revealed no significant difference in PGC induction efficiency for the WT cell line. In contrast, the MUT-NANOS1 cell line displayed a highly significant reduction in efficiency following doxycycline-induced overexpression (****$p < 0.0001$).

13. Levels of ME markers upregulated in MUT-NANOS1 cell lines after overexpression

To assess whether prolonged WNT signaling, supported by upregulated TGF-ß signaling, drives late mesoderm differentiation at the PGC stage, ME marker levels were evaluated (**Fig. 60**). Both MUT cell lines 3.1.8 and 3.2.5.2 showed significant upregulation of LEF1 and TBX3 markers (**Fig. 60**). In contrast, HAND1 and WLS were significantly upregulated in MUT 3.2.5.2, but not in MUT 3.1.8 cell line. Notably, in MUT 3.1.8, HAND1 levels remained unchanged while WLS levels were upregulated; however, high variability among biological replicates prevented statistical significance (**Fig. 60**, top panel), a limitation that additional replicates might have overcome. The consistent upregulation of LEF1 and TBX3 across both MUT-NANOS1 cell lines suggests a shared differentiation trajectory influenced by MUT-NANOS1 overexpression, whereas the variability in other markers points to a more cell line-specific differentiation tendency.

TBX3, a T-box transcription factor, is dynamically expressed during the specification of ME lineages in differentiating mouse ESCs *in vitro* and in developing mouse and *Xenopus* embryos *in vivo* (Weidgang et al., 2013). Forced expression of TBX3 in ESCs promotes ME specification by directly activating key lineage specification factors and indirectly enhancing paracrine NODAL/SMAD2 signalling. For instance, blocking the NODAL/SMAD2 pathway significantly impedes the inductive effects of TBX3, highlighting NODAL as an immediate target of TBX3 actions (Weidgang et al., 2013). Furthermore, TBX3 also 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 signalling (Price et al., 2012).

On the other hand, LEF1, a TCF/LEF TF and nuclear effector of canonical WNT signaling, is essential for mesoderm patterning and cell fate determination (Roël et al., 2009). In *Xenopus*, lef1 expression begins just before mesoderm specification during gastrulation, and its depletion disrupts mesoderm development, including paraxial mesoderm formation and pronephros differentiation (Roël et al., 2009; Molenaar et al., 1998). Similarly, LEF1 enrichment has been observed in human ESC-derived endoderm progenitors, with its downregulation linked to ME marker loss during differentiation (Chu et al., 2016).

Together with TBX3, LEF1 plays a pivotal role in ME lineage specification. For instance, TBX3 promotes ME differentiation by interacting with core signaling pathways, while LEF1 acts as a mediator of WNT signaling for mesoderm patterning. In MUT-NANOS1 overexpressing PGCs, upregulation of TBX3 and LEF1 reflects prolonged activation of WNT and TGF-β pathways, as reflected in the GO analysis of the altered RNA interactome(**Fig. 52A**).

Furthermore, significant HAND1 upregulation in the MUT 3.2.5.2 cell line reinforces its mesoderm differentiation trajectory. For instance, Castillo-Venzor et al. (2023) showed that between 24–32 h of differentiation in activin-induced primed hESCs, a PDGFRA-positive population emerges, representing advanced mesoderm-like cells (aMeLC), concomitant with a decline in nascent

mesendoderm-like cells (MeLC) and emerging MeLC subpopulations. Diffusion pseudotime analysis revealed late upregulation of several advanced mesoderm markers, including HAND1, SNAI2, and GATA6. In the current study, HAND1 upregulation was validated (**Fig. 60**), while SNAI2 and GATA6 transcripts showed significant increase by DESeq analysis following MUT-NANOS1 overexpression (data not shown). As the earliest specified fate, nascent and emergent mesoderm cells show downregulation of pluripotency factors OCT4, SOX2, and NANOG (as validated in **Fig. 56**) and upregulation of several genes that may influence the balance of fates within the EB, including BMP4, WNT5A, CER1, and extracellular matrix genes (Castillo-Venzor et al., 2023).

Notably, BMP4 was significantly upregulated following inducible MUT-NANOS1 expression according to DESeq analysis (data not shown). While BMP4 levels were elevated in both MUT cell lines via RT-qPCR, statistical significance could not be established due to high variability among replicates, particularly in MUT 3.2.5.2 (**Fig. 60**, bottom panel).

On the other hand, the significant upregulation of the WLS marker in MUT 3.2.5.2, along with its upregulation in MUT 3.1.8 despite lacking statistical significance (**Fig. 60**), underscores the importance of WLS. This multipass transmembrane protein is essential for mesoderm induction, acting as a WNT-sorting receptor that transports WNT proteins to the cell surface, a function conserved from cnidarians to humans (Guder et al., 2006). In mice, loss of WLS results in embryonic lethality due to failed mesoderm induction and gastrulation (Fu et al., 2009); its orthologue, Gpr177, shifts from the posterior visceral endoderm and epiblast to the mesoderm at the late-streak stage, paralleling Wnt3 expression, which is vital for axial patterning and PGC specification. In Gpr177 double-knockout embryos, development is arrested at the egg cylinder stage, lacking mesoderm or primitive streak formation (Fu et al., 2009). Similarly, in the context of PGCs in this study, MUT-NANOS1 overexpression leads to increased WLS mRNA levels and enhanced WNT signaling, suggesting that elevated WLS further consolidates mesoderm-like lineage induction in PGCs by promoting WNT secretion and activity.

Lastly, the significant downregulation of WNT5A in both MUT cell lines (**Fig. 60**) was an unexpected result, as WNT5A was reported to be upregulated in late mesoderm lineages (Castillo-Venzor et al., 2023). WNT5A, a key ligand in WNT signaling, plays critical roles in PGC migration and proliferation through the noncanonical WNT pathway mediated by its receptor ROR2. While canonical WNT signaling, involving ligands like WNT3 and WNT3A, is essential for PGC specification and sex differentiation (Aramaki et al., 2013; Tanaka et al., 2013), WNT5A–ROR2 regulates PGC migration via planar cell polarity (PCP) mechanisms, affecting protrusion formation, calcium signaling, and ECM degradation (Slusarski et al., 1997; Nishita et al., 2006; Enomoto et al., 2009; Cantú and Laird, 2017). *In vivo*, WNT5A suppresses β-catenin–dependent canonical WNT signaling through ROR2, as evidenced by increased canonical activity in Ror2 mutants (Mikels et al., 2009). Mutations in Wnt5a result in reduced PGC colonization of the gonadal ridges and disrupted male gonad

development (Chawengsaksophak et al., 2012). Therefore, in PGCs overexpressing MUT-NANOS1, WNT5A downregulation (**Fig. 60**), potentially leads to impaired PGC migration by enhancing canonical WNT signalling, as shown by upregulated β-catenin and TCF complex activity (**Fig. 52A**).

In summary, MUT-NANOS1 overexpression upregulates ME markers, including LEF1, TBX3, HAND1, and WLS, indicating a shift toward ME differentiation driven by prolonged WNT and elevated TGF-β signaling. TBX3 activates lineage-specific factors and enhances NODAL/SMAD2 signaling, while LEF1 mediates essential WNT-driven ME patterning. Increased HAND1 and WLS, which facilitates WNT secretion, further support ME commitment, particularly in the MUT 3.2.5.2 cell line.

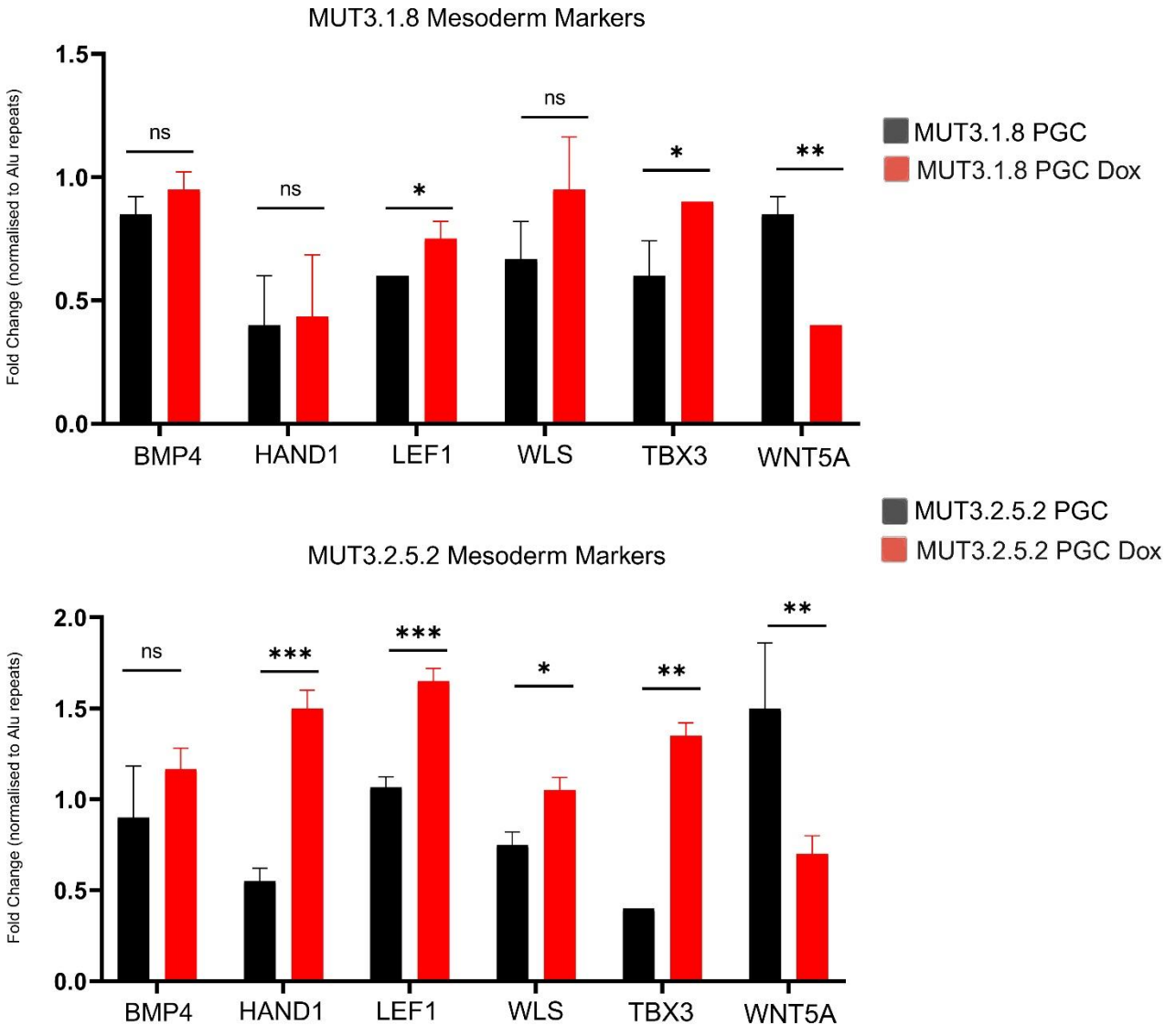


Figure 60. Upregulation of ME markers in MUT-NANOS1 cell lines 3.1.8 and 3.2.5.2 at the PGC stage. Expression levels of the ME markers were measured in MUT-NANOS1 cell lines 3.1.8 (top panel) and 3.2.5.2 (bottom panel) before (black bars) and after (red bars) dox-induced overexpression. Fold changes were normalized to Alu repeats. T-box transcription factor 3 (TBX3) and lymphoid enhancer binding factor 1 (LEF1) are significantly upregulated in both MUT cell lines, whereas heart and neural crest derivatives expressed 1 (HAND1) shows significant upregulation only in MUT 3.2.5.2 (***p < 0.001). Bone morphogenetic protein 4 (BMP4) exhibits an upward trend in both cell lines but does not reach statistical significance. WNT ligand secretion

mediator (WLS), is significantly upregulated only in MUT 3.2.5.2 (*p < 0.05). WNT family member 5A (WNT5A), a non-canonical WNT signalling ligand, is significantly downregulated in both cell lines (**p < 0.01). Statistical analyses were performed using multiple unpaired one-tailed Student's t-tests

14. WNT signalling inhibition enhances PGC markers and downregulates ME markers, reinforcing PGCLC profile following MUT-NANOS1 expression *in vitro*

Lastly, given the persistent elevation of WNT signalling, potentially supported by increased TGF-β signalling (**Fig. 52A**), and the validated upregulation of late mesoderm and ME markers (TBX3, HAND1, LEF1; **Fig. 60**), WNT signalling inhibition was performed to evaluate its potential to reverse the expression of validated PGC and ME markers (**Fig. 56** and **Fig. 60**). For this purpose, the WNT inhibitor Wnt-C59 was applied to PGCs induced to overexpress MUT-NANOS1, specifically in cell line MUT 3.2.5.2 (labelled as MUT_1 in **Fig. 61A–B**). Wnt-C59 mediates WNT inhibition by targeting Porcupine (PORCN), a crucial enzyme required for the palmitoylation, secretion, and activity of all WNT proteins, thereby blocking both canonical and noncanonical WNT signalling pathways (Proffitt et al., 2013).

Following Wnt-C59 treatment, the expression levels of pluripotency and PGC factors OCT4, NANOG, and PRDM14 significantly increased in PGCs overexpressing MUT-NANOS1 compared to those without WNT inhibition (**Fig. 61A**). This is particularly notable, as PRDM14 is a key marker at the late stage of PGC specification (Sybirna et al., 2020), which is the stage when the Wnt-C59 inhibitor was added and overexpression of MUT-NANOS1 occurred. In contrast, MUT-NANOS1 overexpression alone at the PGC stage led to a significant downregulation of OCT4 and NANOG (**Fig. 56**), likely due in part to MUT-NANOS1 binding directly to these transcripts (**Fig. 57C**). By potentially binding to and repressing OCT4 and NANOG, MUT-NANOS1 appears to enhance aberrant WNT signaling, leading to reduced pluripotency and PGC markers and promoting a shift toward mesodermal differentiation in pre-me and PGCLC cells.

The restoration of PGC identity following WNT inhibition is also evidenced by the significant upregulation of NANOS3 (**Fig. 61A**), a key PGC marker, which was previously downregulated in MUT-NANOS1 cell lines and linked to reduced PGCLC differentiation efficiency (**Fig. 56** and **Fig. 59**). The upregulated NANOS3 expression implies potential improvement on PGCLC induction efficiency *in vitro*, counteracting the negative effects of MUT-NANOS1 overexpression. Additionally, WNT inhibition significantly downregulated β-catenin, a key effector of the canonical WNT pathway and a direct binding target of MUT-NANOS1 based on the eCLIP enriched RNA interactome (data not shown). This demonstrates effective suppression of the notably upregulated WNT signalling at the PGC stage (**Fig. 52A**).

Furthermore, WNT inhibition significantly downregulated late mesoderm markers such as LEF1 and HAND1 (**Fig. 61B**), previously upregulated in MUT-NANOS1 cell lines (**Fig. 60**), indicating

a reversion of differentiated cells toward a PGCLC profile. This reversion is marked by an enriched pluripotency network (**Fig. 61A**) and reduced mesoderm progenitor markers (**Fig. 61B**). Similarly, TBX3, a downstream activator of WNT signalling and regulator of the NODAL/SMAD2 pathway (Price et al., 2012; Weidgang et al., 2013), was significantly downregulated (**Fig. 61B**), further suppressing the ME trajectory induced by MUT-NANOS1 expression.

Furthermore, BMP4, which was previously upregulated in MUT-NANOS1 cell lines but not statistically significant (**Fig. 60**), showed a significant reduction after WNT inhibition (**Fig. 61B**). This demonstrates the interplay between BMP and WNT signalling, where WNT inhibition suppresses BMP signalling. BMP4, a TGF-β family signalling protein upregulated in MUT-NANOS1 PGCLCs, plays a dual role in mesoderm induction during gastrulation and in limiting PGC numbers after their specification by promoting differentiation toward mesodermal derivatives like the allantois (Hadas et al., 2024). In mice, BMP4 peaks in the extraembryonic mesoderm (ExM) and gradually declines as chorion and PGC differentiation progress, paralleling the BMP4 downregulation observed after WNT inhibition (**Fig. 61B**). This interplay between BMP4 and WNT signalling is essential for germ-layer specification, as BMP initiates WNT signalling, which activates NODAL, maintaining BMP signalling to drive primitive streak formation (Chhabra et al., 2019). The downregulation of BMP4 following WNT inhibition likely reflects the interruption of WNT-mediated NODAL activation, halting prolonged BMP signaling.

Together, these changes enrich pluripotency markers (OCT4, NANOG) and PGC markers (NANOS3, PRDM14) (**Fig. 61A**) while reducing WNT and mesodermal markers β-catenin (CTNNB1), TBX3, LEF1, HAND1, along with BMP4 (**Fig. 61B**) signaling, thereby potentially restoring the PGCLC trajectory of cells overexpressing MUT-NANOS1.

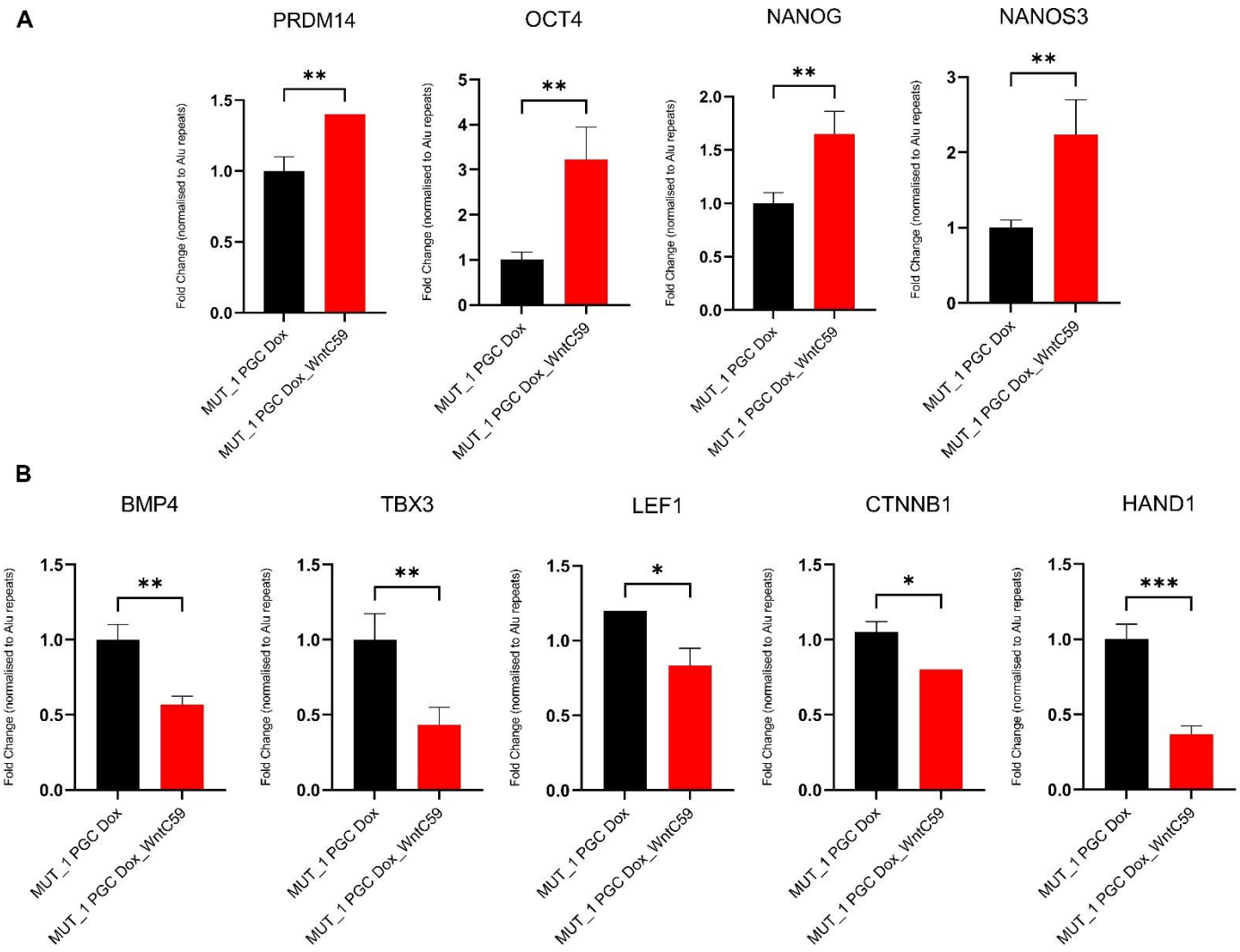


Figure 61. Effects of WNT inhibition (Wnt-C59, 10 µM) on MUT 3.2.5.2 (MUT_1) PGCLCs based on PGC and ME markers levels. A- Treatment with Wnt-C59 significantly restores the expression of PGC markers NANOS3, OCT4, NANOG, and PRDM14 (all markers, $^{**}p < 0.01$), which were previously downregulated in dox-treated PGCs overexpressing MUT-NANOS1. **B-** β-catenin (CTNNB1), a key effector of canonical WNT signalling, is significantly downregulated following Wnt-C59 treatment ($^{*}p < 0.05$), confirming effective inhibition of WNT signaling. BMP4, a downstream target of BMP signalling, is also significantly downregulated ($^{**}p < 0.01$). Mesoderm markers TBX3, LEF1, and HAND1, which were upregulated in dox-induced MUT 3.2.5.2 PGCs, are significantly downregulated following WNT inhibition ($^{*}p < 0.05$; $^{**}p < 0.01$; $^{**}p < 0.001$). Fold changes are normalized to Alu repeats, and statistical analyses were performed using unpaired one-tailed Student's t-tests.

DISCUSSION

MUT-NANOS1 exhibits GOF characteristics at the early stage of human PGC specification
Combined eCLIP and RNA-seq data show that MUT-NANOS1 regulates a much larger number of mRNA targets (148) compared to WT-NANOS1 (19), with only 15 targets common between them at the pre-me stage. At the PGC stage, MUT-NANOS1 regulates 385 targets, whereas the WT counterpart regulates only 6, with just 3 targets overlapping. Moreover, eCLIP analysis revealed that WT-NANOS1 predominantly targets the CDS and, to a moderate extent, the 3′UTR, with relatively sparse intronic binding after normalization against the lengths of introns, exons, and both the 5′UTR and 3′UTR, using the 5'UTR as reference (**Fig. 43E**). In contrast, MUT-NANOS1 exhibits a distinct binding profile, showing similar enrichment in the CDS and 3′UTR along with additional binding to the 5′UTR (**Fig. 43E**). This shift, particularly the marked enrichment at the 3′UTR, suggests that MUT-NANOS1 may inappropriately acquire a repressive role by binding to 3′UTRs where it normally would not, leading to excessive repression of mRNAs essential for proper PGC differentiation. Since binding in these regions is fundamental to post-transcriptional gene regulation by influencing mRNA stability, localization, and translation, these findings indicate a possible dominant gain-of-function effect induced by the p.[Pro34Thr; Ser78del] mutation.

Dominant GOF mutations are characterized by milder structural impacts and tend to be enriched at protein interfaces, often clustering in three-dimensional space rather than causing widespread structural changes (Gerasimavicius et al., 2022). However, our initial analysis comparing the structure of MUT-NANOS1 to its WT counterpart using a molecular dynamics simulation approach did not conclusively support this rule (not shown).

The double mutation p.[Pro34Thr; Ser78del] of NANOS1 exhibits additional GOF characteristics. Both mutations are located in the N-terminal IDR (**Introduction Fig. 4**), where they can facilitate aberrant interactions or activation states that are not typical of the WT protein. This observation is consistent with the idea that GOF mutations in less ordered regions establish novel interactions rather than altering structured domains (Li et al., 2018). Notably, the double mutation encompasses the NIM functional motif, which is important for recruiting the CNOT deadenylation complex for repression. Furthermore, dominant GOF mutations often cluster near functional sites, inducing aberrant activity without markedly disrupting overall protein structure (Sivley et al., 2018). While direct binding of individual targets requires validation, these findings suggest that the acquired post-transcriptional repressive role of MUT-NANOS1 persists from the pre-me stage into the PGC stage, likely impairing several pathways required for proper PGC development, as discussed below.

MUT-NANOS1 mediates repression in pre-me by downregulating cell cycle phase transition and mitotic cell cycle pathways The acquired premature repression of MUT-NANOS1 was initially placed on the cell proliferation pathways, following initial observations by Illaslan et al. (2022) that WT-NANOS1 overexpression in the TCam-2 human germ cell line model led to an accumulation of cells in the G1 and G2/M phases, while MUT-NANOS1 produced the opposite effect. Consistent with these findings, GSEA analysis (**Fig. 44** and **Fig. 45**) in this study revealed a downregulation of cell cycle-related target mRNAs in pre-me cells overexpressing MUT-NANOS1, suggesting a disruption of cell cycle–dependent differentiation signals. In addition, eCLIP data indicates that MUT-NANOS1 may further influence the cell cycle by binding to the pluripotency factor MYC (at its CDS) and the cell cycle regulator cyclin D1 (at its 3'UTR), both of which are downregulated in pre-me cells (**Fig. 49**). MYC primarily functions as a transcription factor that promotes cell proliferation (Dang et al., 2006), while cyclin D1 regulates the G1 phase of the cell cycle, with MYC also influencing the S phase by enhancing cyclin D1 expression (Lecarpentier et al., 2019).

Overall, pluripotency is linked to cell-cycle regulation, as studies in mouse ESCs have shown that their pluripotent state is associated with a shortened G1 phase and a lack of G1 checkpoint regulation (Coronado et al., 2013).

Since early PGC specification relies on precisely coordinated cell cycle progression and signalling pathways, the findings in **Fig. 44** and **Fig. 45** support Pauklin and Vallier's (2013) observation that premature cell cycle repression by MUT-NANOS1 may prevent cells from entering the proper differentiation window. By directly binding to and repressing key cell cycle regulators MYC and cyclin D1, as evidenced by eCLIP data and confirmed in **Fig. 49**, MUT-NANOS1 disrupts proliferation and downstream pathways that are crucial for establishing PGC competence during the pre-me transition (Tang et al., 2022). Consequently, many cells may fail to progress synchronously through the necessary cell cycle stages, ultimately hindering PGC formation.

Suppression of pluripotency factors OCT4 and NANOG by direct binding of MUT-NANOS1 Based on eCLIP data, MUT-NANOS1, unlike the WT counterpart, binds to the pluripotency factors OCT4 and NANOG, specifically at the 5'UTR of NANOG mRNA and the 3'UTR of OCT4 mRNA (**Fig. 57C**). This binding likely mediates the initial suppression of OCT4 and NANOG during the pre-me stage (**Fig. 57A-B**) and sustains their reduced expression in PGCs (**Fig. 56**), where they play more prominent roles, as confirmed by time course analysis reflecting their endogenous expression levels (**Fig. 57D-E**). In turn, this downregulation disrupts the regulatory network essential for PGC competence in pre-me, compounded by reduced WNT signalling, and impairs the PGC differentiation trajectory when MUT-NANOS1 is overexpressed on Day 3 of a 4-day differentiation protocol.

Notably, OCT4 and NANOG are pivotal TFs essential for the establishment and maintenance of pluripotency during early embryogenesis (Irie et al., 2014). Their coordinated expression is also

crucial for germ cell competence and differentiation, guiding the specification of PGCs *in vivo* and their *in vitro* differentiation into PGCLCs. The expression of OCT4 and NANOG is maintained through upstream binding of SOX17 and TFAP2C to their enhancers located upstream of their promoter sequences. In turn, NANOG and OCT4 further cooperate with SOX17 and TFAP2C, along with epigenetic regulators, to remodel chromatin toward a human PGC fate (Tang et al., 2022).

Given the critical roles of OCT4 and NANOG in maintaining pluripotency and initiating germ cell-specific gene expression, their diminished expression likely compromises the transition from pre-me to PGCs and the maintenance of the germ cell profile following cytokine-induced differentiation.

WNT signalling pathway is significantly downregulated in the MUT-NANOS1-enriched interactome of pre-me cells Following an overwhelming enrichment of downregulated pathways post-GSEA (illustrated in the UpSet plot, **Fig. 44**), GO analysis was performed on the downregulated and enriched mRNAs of MUT-NANOS1 (**Fig. 46**). Notably, the analysis reveals a marked reduction in WNT signalling (**Fig. 46**), a pathway fundamental to embryonic development and PGC lineage specification. The statistically significant repression of WNT pathway components bound by MUT-NANOS1 likely disturbs PGC formation and diminishes the overall differentiation capability of pre-me cells.

Analysis of the enriched and altered MUT-NANOS1 RNA interactome revealed downregulation of several mRNAs associated primarily with canonical WNT (**Fig. 47**). This pathway is known to play critical roles in cell fate determination, proliferation, and survival (Mohammed et al., 2016). In the context of PGC specification, WNT3, a key canonical WNT pathway component expressed in the human epiblast at E5.5, primes cells for BMP4 responsiveness and PGC induction. *In vitro*, during the pre-me stage, activin A further enhances PGC potential by upregulating OCT4, NANOG, NODAL, and WNT3 while suppressing BMP signaling. This transient conversion before differentiation boosts human PGC formation, underscoring canonical WNT3's critical role in germ cell induction (Cheng et al., 2022).

Frizzled receptor 10 (FZD10) (**Fig. 47** and **Fig. 48A-B**), normally an activator of canonical WNT signalling, is significantly repressed. This receptor has been shown to work synergistically with select classes of WNT ligands in upregulating canonical signalling (Terasaki et al., 2002). Similarly, secreted Frizzled-related protein 2 (SFRP2), a WNT modulator, (**Fig. 48A-B**) is downregulated following MUT-NANOS1 overexpression and binding as indicated by eCLIP data. In PGC-competent pre-me progenitor cells, SFRP2-mediated WNT interactions are essential for regulating PGC competence (Castillo-Venzor et al., 2023). Typically, SFRP2 antagonizes WNT by competing for FZD binding against WNT morphogens, but its effect can be context-dependent, such as promoting canonical

WNT/ß-catenin signalling in cardiac fibroblasts, evidenced by increased expression of AXIN2 and WNT3A, as well as the accumulation of nuclear ß-catenin (Lin et al., 2016).

TCF7L1, another significant WNT component (**Fig. 48A-B**), is notably downregulated following MUT-NANOS1 binding (**Fig. 48A-B**). As a member of the TCF/LEF family of high mobility group (HMG)-box DNA-binding proteins, TCF7L1 primarily represses WNT targets or can act as a weak activator (Sierra et al., 2018). Unlike TCF7 and LEF1 nuclear factors, which are strong activators, TCF7L1 is the most highly expressed TCF/LEF factor in undifferentiated primed hESCs, where it suppresses genes associated with gastrulation and primitive streak formation (Nodal, WNT3, and BMP4) to maintain hESC pluripotency (Sierra et al., 2018). Moreover, Sierra et al. (2018) found that TCF7L1 is less integrated with the core pluripotency TFs (OCT4 and NANOG) and is downregulated by BMP4, a known inducer of ME. In contrast, at the pre-me stage (6 h post Nodal and WNT activation), the cells have not been treated with exogenous BMP to downregulate TCF7L1, therefore, the premature repression induced by MUT-NANOS1 could affect the pluripotency of the stem cell population and the regulation of WNT signalling, which is crucial for the cells to respond appropriately to differentiation cues as precursors to PGCLCs.

TCF4 (TCF7L2) similarly exhibits downregulation (**Fig. 48A-B**), possibly mediated by MUT-NANOS1 eCLIP-enriched binding. This core nuclear transcription factor in the WNT signalling pathway exhibits dual functions in transcriptional activation and repression (Salinas et al., 2008). In the absence of WNT signals, TCF4 binds to WNT target gene promoters to inhibit their expression. Upon WNT pathway activation, β-catenin accumulates, enters the nucleus, and forms a complex with TCF4, leading to the activation of WNT target genes (Lee et al., 1999). Therefore, the untimely downregulation of TCF4 likely impairs WNT-driven transcriptional regulation.

Collectively, the downregulation of these WNT pathway components, ranging from receptors and modulators (FZD10, SFRP2) to nuclear transcription factors (TCF7L1, TCF4), suggests a premature suppression by the MUT-NANOS1 p.[Pro34Thr; Ser78del] variant likely mediated by direct binding, which would require further validation via luciferase assays. This untimely inhibition compromises canonical WNT activation critical for primitive streak-like induction and Nodal signalling in the pre-me stage, potentially undermining proper PGC specification.

Furthermore, Cyclin D1, which is downregulated in pre-me cells (**Fig. 47** and **Fig. 49A**), not only manages cell cycle transition but also serves as a canonical WNT target regulated by TCF/LEF transcription factors through a highly conserved binding element near its promoter region. MUT-NANOS1 binding to the 3'UTR of cyclin D1 (eCLIP data), likely destabilises the transcript leading to its downregulation. *In vitro* studies have shown that degradation-resistant β-catenin mutants enhance cyclin D1 expression and promoter activity, highlighting β-catenin's positive regulation of cyclin D1 (Shtutman et al., 1999).

Likewise, decreased mRNA levels of the surface marker CD24 (**Fig. 47** and **Fig. 49B**) reflect another WNT target gene showing diminished transcript abundance in pre-me cells. Collier et al. (2017) found that CD24 transcripts are more abundant in post-implantation epiblast cells, supporting the classification of CD24 as a primed-state marker. Because pre-me precursors are derived from hESCs that represent post-implantation epiblast cells, existing in a state between primed and naïve pluripotency, leaning toward the primed progenitor lineage (Chen et al., 2019), the presence of this surface marker is expected at high levels at the pre-me stage. Its downregulation implies that MUT-NANOS1 disrupts key developmental signatures of primed epiblast-like cells, potentially interfering with the transition toward the PGC lineage. Lastly, AXIN2 and MYC (c-MYC), both canonical WNT targets (Rennoll et al., 2015), are also downregulated (**Fig. 49C–D**). AXIN2 operates within the destruction complex and participates in a negative feedback loop that modulates the duration of the canonical WNT response. Its WNT-responsive enhancer elements (WREs) lie in the 5′ promoter region and downstream of the transcription start site (Leung et al., 2002). Canonical WNT signalling typically induces AXIN2 expression, thus initiating this negative feedback mechanism. According to eCLIP data, MUT-NANOS1 may bind at the 3′UTR of AXIN2, suggesting an impact on its mRNA stability.

Furthermore, the WNT pathway governs hESC behaviour by either promoting self-renewal or inducing differentiation, depending on the cellular localization of β-catenin. Stabilization of β-catenin in the cytoplasm, mediated in part by AXIN2, supports the hESC pluripotent identity. Conversely, when β-catenin translocates into the nucleus and interacts with TCF transcription factors, it drives differentiation (Kim et al., 2013). In this study, both cytosolic and nuclear WNT components are disrupted (**Fig. 47** and **Fig. 48A–B**), affecting pluripotency and differentiation.

During the cell cycle, β-catenin expression oscillates, peaking at the G2/M transition. Its downstream target, AXIN2, similarly reaches maximum expression at this stage (Hadjihannas et al., 2012). Since β-catenin may also be bound by MUT-NANOS1 (eCLIP data), premature repression of these interactions could further diminish WNT signalling.

Lastly, MYC, previously noted for its role in cell proliferation, dual function in the G1 phase, and influence on cyclin D1, is also a target of the WNT signalling pathway. Its expression is regulated by WREs located both near gene boundaries and several hundred kilobases away from the transcription start site (Yochum et al., 2008). Elevated nuclear β-catenin drives the activation of numerous genes, most notably cyclin D1 and MYC (Lecarpentier et al., 2019). Although MYC represses differentiation-related genes during iPSC formation (Wey and Knoepfler, 2010), maintaining pluripotency by preserving high levels of cellular metabolism, including protein and DNA synthesis (van Riggelen et al., 2010; Dominguez-Sola et al., 2008), MUT-NANOS1-mediated repression reduces these pro-pluripotency and pro-proliferation effects (**Fig. 44**).

These findings highlight multiple layers of WNT regulation during embryogenesis. Initially, WNT signalling drives primitive streak cells toward ME progenitors. Subsequently, it collaborates with BMP

and Nodal pathways to direct the anterior/posterior fate of mesoderm derivatives (Hernández-Martínez et al., 2024). The interplay of WNT components such as AXIN2 and MYC is crucial for controlling cell proliferation and differentiation. Based on eCLIP data, MUT-NANOS1 likely binds these mRNAs, influencing their stability and translation. Consequently, both pluripotency factors (e.g., MYC) and differentiation factors (e.g., AXIN2) are downregulated, emphasizing MUT-NANOS1's significant impact on WNT signalling and its broader role in undermining normal cell fate decisions toward PGCLCs.

Potential mechanism of suppression of MYC, SFRP2 and CDH2 mRNA levels by MUT-NANOS1 directed by specific motif sequence within their CDS' In mice, Nanos2 binds RNA directly via the consensus sequence AUKAAWU (with each zinc-finger recognizing three nucleotides; Choo and Klug, 1994; Codino et al., 2021) at the 3′UTRs of its target mRNAs—a mechanism distinct from Drosophila Nanos, which requires Pumilio for motif-specific binding (Sonoda and Wharton, 1999; Weidmann et al., 2016).

This coordinated, motif-based regulation of gene expression highlights the need to investigate whether NANOS1 in human PGCs employs a similar mechanism. To address this, motif analysis of suppressed WNT pathway–related target mRNAs identified the SAAGAA (identified as GAAGAA) motif in 16.87% of the sequences examined, with this motif present in three validated targets, MYC (**Fig. 49C**), SFRP2 (**Fig. 48A–B**), and CDH2 (**Fig. 50H**), specifically within their CDS' (**Fig. 50A**). Although further luciferase assays are needed to confirm that MUT-NANOS1 binds these targets and cooperatively downregulates them, it is noteworthy that while the CDS ranks third in enrichment in MUT-NANOS1's altered RNA interactome (after the 3′UTR and introns; **Fig. 43D**), normalization to each region's genomic length shows the CDS as the most enriched site for WT-NANOS1 and comparably enriched alongside the 3′UTR in MUT-NANOS1 (**Fig. 43E**).

Since the CDS encodes protein sequences, changes in NANOS1 binding could affect translation efficiency, mRNA stability, or other downstream regulatory events, underscoring the importance of motif-based analysis in elucidating NANOS1's role in human PGCs.

While MYC and SFRP2 have notable roles in cell proliferation, cell cycle transition, and context-dependent WNT pathway modulation, CDH2 emerges as another transcript negatively affected by MUT-NANOS1 overexpression in pre-me cells. Typically, CDH2 is upregulated during gastrulation-associated epithelial-mesenchymal transition (EMT) by nascent mesoderm cells emerging from the primitive streak, along with T expression, generating a posterior stripe of CDH2 and T-expressing cells by E6.5 in mice (Hernández-Martínez et al., 2019).

Within this framework, MUT-NANOS1 disrupts the WNT pathway at multiple levels, not only by modulating cell–cell junctions through repression of CDH2 transcript expression, but also by downregulating SFRP2 and MYC, potentially via its binding to the GAAGAA motif.

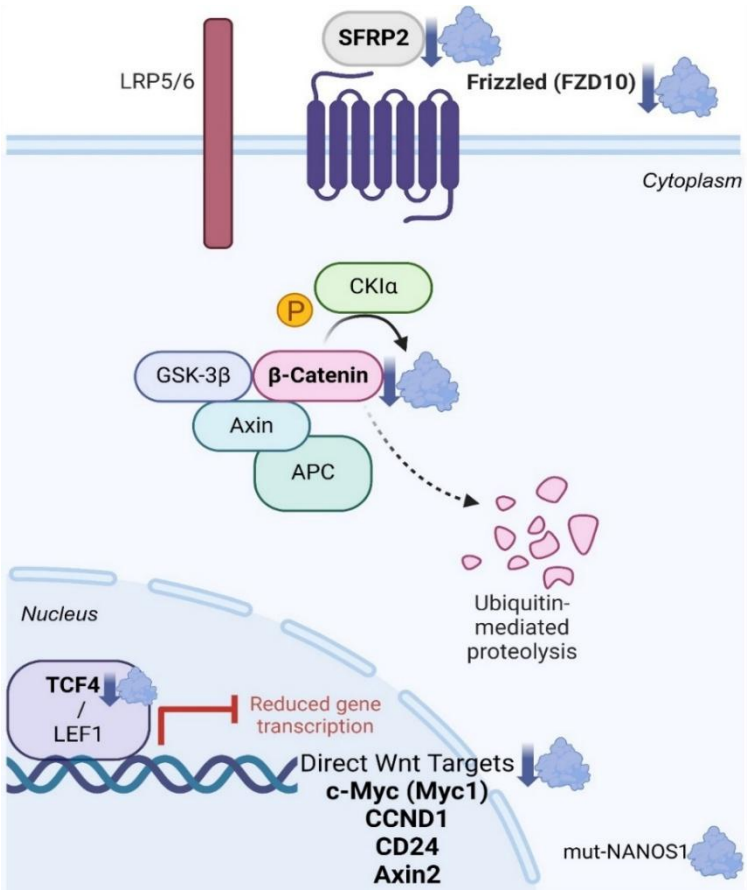
Notably, both SFRP2 and MYC show high expression at the pre-me and stem cell stages (**Fig. 50C-E**), aligning with WNT/β-catenin's dual function in hESCs. In one mode, it maintains pluripotency via cytoplasmic AXIN2 stabilization, and in another, it drives differentiation when β-catenin enters the nucleus and partners with TCF/LEFs (Kim et al., 2013). This functional versatility is well-documented in heterogeneous hESC cultures (Ávila-González et al., 2021). Thus, the downregulation of MYC, SFRP2, and CDH2, likely through MUT-NANOS1 binding at the GAAGAA motif within their CDS (**Fig. 50A**), may compromise critical processes such as pluripotency maintenance, cell division, cell adhesion, and signalling modulation.

Importantly, these targets were not enriched in the WT-NANOS1 RNA interactome, even though the CDS is also the most enriched binding site for WT-NANOS1 following normalisation of eCLIP data (**Fig. 43E**). This suggests that WT-NANOS1 did not employ its canonical repression, mostly mediated via 3'UTR, to regulate the expression of potential target RNAs in this stage of stem cell differentiation. Consequently, MUT-NANOS1 appears to exert a GOF repressive mechanism, illustrated by its unique motif recognition and target binding, that diminishes WNT signalling and may hinder the ability of pre-me precursors to efficiently progress toward ME during which competent sub-clusters give rise to PGCLCs.

WNT pathway downregulation by MUT-NANOS1 reduces pre-me cells' competence to PGCs, evidenced by lower EOMES expression To evaluate the impact of WNT downregulation on PGC competence in pre-me cells, EOMES, a known PGC competence marker and target of WNT signalling, showed significant downregulation (**Fig. 51A–B**) following MUT-NANOS1 overexpression, further highlighting the impact of the GOF premature suppression of the WNT pathway by MUT-NANOS1. Studies have demonstrated that EOMES is a key regulator of PGC competence in pre-me cells, where an increase in its levels enhances human PGC competence (Chen et al., 2017); however, further activation of EOMES alongside mesoderm factors can hinder PGC specification.

In summary, these findings indicate that MUT-NANOS1 overexpression suppresses WNT signalling, as evidenced by the marked downregulation of EOMES, an essential regulator of PGC competence and downstream target of WNT (Chen et al., 2017; Castillo-Venzor et al., 2023). Because proper EOMES expression is crucial for transitioning pre-me cells into PGCs *in vitro*, its reduction under MUT-NANOS1 overexpression likely compromises the specification and subsequent differentiation of PGCs *in vivo*. This underscores the delicate balance of WNT signalling, EOMES

regulation, and other early developmental factors required for effective human PGC formation (repressed components are highlighted in **Discussion Fig. 1**).



MUT-NANOS1 negatively disrupts PGCLC differentiation by upregulating WNT and TGF-ß signalling pathways Following GO analysis of the MUT-NANOS1 altered RNA interactome at the PGC stage, the less statistically upregulated pathways compared to downregulated processes (**Fig. 52A**) predominantly involve structural components, signalling, and morphogenetic processes indicative of differentiation and adaptation. By comparison, the downregulated pathways reflect reduced responses for correct protein folding, stress response, and cell adhesion (**Fig. 52A-B**). These observations point to a trade-off wherein the cell prioritizes structural and differentiation-related processes over induced cellular stress management and adhesion stability. Notably, the reduction in adhesion could be associated with the upregulated WNT and TGF-β signalling pathways highlighted in the GO analysis (**Fig. 52A**). These pathways likely induce extracellular matrix remodelling toward somatic lineages and promote cellular reorganization, despite expectations that PGCs at this stage should maintain their germline identity.

Previous research on human PGC differentiation *in vitro* has identified distinct roles and timing requirements for BMP, WNT, and Nodal (TGF-beta superfamily member) signaling pathways.

Continuous BMP signaling is essential throughout the first two days (48 hours) of differentiation, whereas WNT signaling is only required during a brief window, specifically, within the first 24 hours during which BMP2/4 is added to the culture medium (Jo et al., 2022).

In the present study, however, during the pre-me stage (**Fig. 46**), the WNT pathway was found to be downregulated following MUT-NANOS1 overexpression, potentially due to the acquired protein's repressive role following its binding to crucial components of the pathway, as enriched in the eCLIP data (**Fig. 47**).

By contrast, the PGC stage comprises differentiated PGCLCs at Day 4 following exogenous BMP4 treatment, which includes early WNT activation via a GSK3β inhibitor and simultaneous Nodal activation via Activin A. In this setting, MUT-NANOS1 overexpression was induced on Day 3 of differentiation, enabling the cells to follow the standard trajectory toward PGCs via BMP, WNT, and Nodal signalling. Typically, these pathways are downregulated by Day 4 of differentiation as PGC specification is established. However, prior to this point, the primary function of WNT is to induce Nodal expression (Jo et al., 2022), a TGF-beta superfamily member involved in crucial embryonic processes such as mesoderm formation, anterior patterning, and left-right axis specification (Kuehn et al., 2001).

Nodal activation of the TGF-beta pathway is indispensable for human PGC induction; notably, external Nodal stimulation alone can rescue human PGC induction *in vitro* even when WNT signalling is inhibited (Jo et al., 2022). In the PGC-stage enriched RNA interactome, MUT-NANOS1 overexpression leads to upregulation of both WNT signalling (via the ß-catenin-TCF complex) and the TGF-beta pathway (**Fig. 52A**). Precise timing and duration of Nodal signalling are critical for determining cell fate, as shorter exposure to Nodal favours PGC differentiation, whereas prolonged high levels of Nodal and subsequent TGF-β pathway upregulation drive endodermal differentiation (Jo et al., 2022). This shift may explain why differentiated cells display altered trajectory fates following MUT-NANOS1 overexpression, despite normally being committed to the PGC trajectory by Day 4 of *in vitro* differentiation.

Under normal conditions, BMP and WNT signalling activate Nodal with a delay, peaking around 24 h (Heemskerk et al., 2019; Chhabra et al., 2019). In contrast, a high dose of exogenous Activin, mimicking Nodal, rapidly suppresses TFAP2C and induces FOXA2 within 42 h, precluding germ cell fate commitment. In the current system, PGCs have already been exposed to elevated Nodal levels; thus, additional TGF-β (**Fig. 52A**) may further suppress TFAP2C and SOX17, diverting cells toward endodermal pathways. Meanwhile, increased WNT signalling could push a mesodermal trajectory. Ultimately, overexpressing MUT-NANOS1 on Day 3 disrupts the precise timing of WNT activation, potentially reactivating BMP and Nodal (via TGF-ß activation) pathways normally inactivated post-PGC commitment.

Upregulation of WNT components in the PGC Stage: Overcoming insufficient MUT-NANOS1 mediated post-transcriptional suppression To further investigate the upregulated WNT pathway components, specific targets were examined and validated. Among those bound by MUT-NANOS1, LRP5 was significantly upregulated at the PGC stage (**Fig. 54A**), despite its predominant role in somatic populations (**Fig. 54B**). In contrast, LRP5 mRNA remained mostly unchanged at the pre-me stage, indicating that MUT-NANOS1 may modulate WNT components in a stage-dependent manner. LRP5, together with its co-receptor LRP6, plays a pivotal role in the WNT/β-catenin signalling pathway by forming a receptor complex with Frizzled receptors on the cell membrane. This complex, along with intracellular components such as Dishevelled (DvL), AXIN, APC, GSK-3β, CK-1α, β-catenin, and TCF/LEF transcription factors, initiates a signalling cascade essential for cell fate determination during both germline and somatic development (Liu et al., 2022). The intracellular domains of LRP5/6 contain phosphorylation sites that, when activated, trigger the canonical WNT/β-catenin pathway (Niehrs and Shen, 2010). In the absence of WNT ligands, β-catenin is targeted for degradation by the "destruction complex" comprising APC, AXIN, CK1, and GSK3, thereby preventing the transcription of WNT-responsive genes. Upon WNT ligand binding, however, the LRP5/6-FZD receptor complex sequesters the destruction complex, allowing β-catenin to accumulate and translocate to the nucleus to activate genes critical for cellular processes (Reyes et al., 2020). In human cells, upregulation of LRP5 and LRP6 has been linked to differentiation processes. For instance, during BMP2-induced differentiation of pluripotent marrow stromal cells along the osteoblastic lineage, elevated expression of these receptors enhances canonical WNT signalling, supporting both the proliferation of pre-osteoblasts and terminal osteoblast differentiation (Gong et al., 2001).

In this study, the upregulation of LRP5 in the PGC stage indicates an enhancement of canonical WNT signalling. This increase may represent an adaptive response to altered WNT activity, potentially supported by the upregulation of additional signaling pathways, such as TGF-β.

Conversely, LRP4, another member of the low-density lipoprotein receptor family, bound by MUT-NANOS1 (eCLIP data), is suppressed at the PGC stage. Notably, LRP4 mRNA levels, which are initially downregulated in pre-me cells, undergo further significant reduction in the PGC stage. This finding is supported by RT-qPCR validations in independently generated MUT-NANOS1 cell lines (**Fig. 54C**), which also aligned with time-course analyses indicating that LRP4 plays a prominent role during the PGC stage prior to MUT-NANOS1 overexpression (**Fig. 54D**).

Unlike LRP5, LRP4 functions as an inhibitor of WNT signalling despite its structural similarities to LRP5/6, because its intracellular domain lacks the motifs necessary for WNT co-receptor activity (Willnow et al., 2012). Overexpression studies have demonstrated that LRP4 reduces canonical WNT signalling activity, underscoring its antagonistic effect on LRP5/6 (Johnson et al., 2005; Li et al., 2010). Additionally, LRP4 interacts with WNT antagonists such as DKK (Dickkopf-related protein 1)

and WISE (also known as SOSTDC-1 or Sclerostin domain-containing protein 1) via its extracellular domain, thereby modulating WNT/β-catenin signalling by inhibiting LRP5/6-mediated activation.

Collectively, these findings indicate that MUT-NANOS1 continues to suppress certain WNT-associated components at the PGC stage. However, additional influences, such as TGF-β signalling and potentially insufficient repression by MUT-NANOS1, may further stabilize and augment canonical WNT signalling, especially via LRP5. This enhanced WNT activity may, in turn, reinforce the downregulation of LRP4, prematurely repressed in pre-me by MUT-NANOS1, leading to prolonged β-catenin nuclear accumulation and sustained TCF/LEF-driven gene activation. The potential binding of MUT-NANOS1 to both LRP4 and LRP5 mRNAs, alongside their antagonistic interplay, highlights the importance of receptor-level regulation in shaping WNT signalling and determining developmental outcomes.

Insufficient MUT-NANOS1-mediated repression of BCL9L mRNA drives crosstalk between upregulated WNT and TGF-β signalling pathways To further investigate WNT and TGF-β crosstalk, BCL9L mRNA levels were examined, as a target potentially bound by MUT-NANOS1 and as a key co-activator of the WNT pathway. GO analysis revealed that BCL9L is associated with TGF-β signalling and is upregulated similarly to LRP5 (**Fig. 53**). Although BCL9L expression increased similarly to LRP5, this upregulation reached statistical significance only when compared to the unedited W15 line rather than the uninduced MUT-NANOS1 controls (**Fig. 54E**). This discrepancy suggests either a cell line–specific effect or that MUT-NANOS1 binding does not adequately repress BCL9L, allowing other pathways to drive its activation.

BCL9L plays a critical role in mediating β-catenin's nuclear activity and influences cellular adhesion, particularly during epithelial-mesenchymal transition (EMT). Activation of the WNT pathway leads to cytoplasmic accumulation of β-catenin, its subsequent nuclear translocation, and interaction with co-factors such as BCL9 and BCL9L to drive transcription of WNT target genes that promote proliferation and migration (Ring et al., 2014). In pancreatic cancer models, BCL9L modulates EMT by affecting E-cadherin (CDH1) levels and β-catenin localization; knockdown of BCL9L increases membrane-associated E-cadherin, sequestering β-catenin and delaying TGF-β–induced EMT, whereas its overexpression enhances β-catenin's nuclear translocation, thereby promoting its transcriptional activity (Sannino et al., 2016). External stimuli such as EGF and TGF-β further induce EMT by downregulating epithelial markers and upregulating mesenchymal markers (Thiery et al., 2009).

These observations suggest that BCL9L upregulation following MUT-NANOS1 overexpression, likely in response to activation via TGF-β signalling, shifts the balance of β-catenin's functions toward nuclear transcription, which is crucial for cell behaviour during EMT. The interplay between TGF-β signalling and BCL9L underscores a complex regulatory relationship, positioning BCL9L as a potential nexus for crosstalk between the WNT and TGF-β signalling pathways. Moreover,

GO analysis on the altered and enriched RNA interactome of MUT-NANOS1 in PGCs reveals a downregulation of cell adhesion processes, notably cadherin binding (**Fig. 52B**), further corroborating the effect of upregulated BCL9L mRNA levels.

Overall, increased BCL9L levels indicate that MUT-NANOS1 overexpression, combined with incomplete repression of this target mRNA, disrupts PGC identity by misregulating the canonical WNT pathway and associated TGF-β signalling (**Fig. 52A**). This dysregulation can promote EMT and mesoderm-like differentiation.

Notably, although these WNT related components were identified from the altered MUT-NANOS1 RNA interactome in PGCs, the data rely on pre-me stage eCLIP analysis due to insufficient *in vitro* PGC numbers for this analysis. At the PGC stage, these transcripts may no longer be directly bound by MUT-NANOS1 but instead reflect indirect consequences of MUT-NANOS1's repressive effects on other regulatory targets crucial for PGC identity, or insufficient repression to counter TGF-β–driven WNT activation.

MUT-NANOS1 mediates EMT via activation of the TGF-β pathway and suppression of E-cadherin in PGCs In addition to WNT signalling, TGF-ß pathway was highlighted as one of the pathways comprising of target mRNAs upregulated in the altered, eCLIP enriched RNA interactome of MUT-NANOS1 (**Fig. 52A**). This observation aligns with recent findings by Rosemann et al. (2024), who demonstrated that NANOS1 acts as a critical negative regulator of TGF-ß signalling in an oral cancer model. In their study, TGF-ß signalling was shown to mediate downregulation of NANOS1 expression as one of the most downregulated RBPs. Furthermore, functional assays showed that overexpression of NANOS1 reduced cancer cell migration and invasion by suppressing EMT induced by TGF-ß signalling. Mechanistically, NANOS1 binds to TGFBR1 mRNA, promoting its degradation and thereby inhibiting downstream TGF-ß/SMAD signalling. This suggests that NANOS1 serves as a key negative regulator of TGF-ß signalling, with potential implications for controlling cancer progression and metastasis (Rosemann et al., 2024).

In the current study, TGFBR1 mRNA is not enriched in the RNA interactome of WT-NANOS1 but is bound by the MUT counterpart at the 3'UTR (eCLIP data). RNA sequencing and DESeq analysis comparing WT and MUT-NANOS1 PGCs following overexpression revealed that TGFBR1 mRNA levels were significantly upregulated (p. adjust = 0.0017; 1.36-fold increase) in the MUT-NANOS1 cell line, which is consistent with GO analysis highlighting increased TGF-β pathway activity (**Fig. S17**). This mirrors the findings from the altered RNA bound interactome for MUT-NANOS1, where upregulated mRNA targets involved in TGF-ß signalling were observed (**Fig. 52A**). Notably, when comparing MUT-NANOS1 PGCs to their uninduced controls, TGFBR1 mRNA levels were not significantly altered, and consequently, this mRNA was not represented in the altered RNA interactome. This suggests that the repression expected from MUT-NANOS1 following potential binding of the

TGFBR1 mRNA was insufficient and not significant enough, leading to enhanced TGF-ß signalling. In turn, this elevated pathway correlated with a significant downregulation of endogenous NANOS1 mRNA levels by 3.3-fold (p. adjust = 4.52e-5) observed in the MUT-NANOS1 PGCs following DGE against the uninduced control cells (RNA-seq. data), a change not observed in WT-NANOS1 cell line.

These results indicate that, despite its potential binding to TGFBR1 mRNA, MUT-NANOS1 fails to repress it effectively, thereby facilitating downregulation of endogenous NANOS1 due to activated TGF-ß signalling as was noted by Rosemann et al. (2024), thereby facilitating upregulation of EMT consequently through downstream effects such as downregulation of E-cadherin (CDH1). Consistent with this mechanism, CDH1 expression was validated in two MUT cell lines, where a marked downregulation was observed (**Fig. 55A**), thereby supporting the expected outcome of the insufficient repression of TGF-ß potentially mediated by MUT-NANOS1.

Notably, GO analysis in the PGC model following MUT-NANOS1 overexpression further revealed significant downregulation of terms related to “cadherin binding,” “focal adhesion,” and “cell adhesion molecule binding” (**Fig. 52B**). CDH1 not only plays a central role in maintaining epithelial integrity but also serves as a link between the TGF-β and WNT pathways via interactions with β-catenin which binds to both cadherins and TCF/LEF transcription factors, playing a crucial role in stabilizing E-cadherin (CDH1) at the cell surface and anchoring adherens junctions (AJs) to the cytoskeleton (Huber et al., 2001; Yap et al., 1997). During EMT, the loss of CDH1 disrupts these junctions, while the gain of CDH2 may transiently increase available CTNNB1 for signalling (Wal & van Amerongen, 2020), potentially reflecting the upregulated "beta-catenin-TCF nuclear complex" term identified in the GO analysis of altered RNA interactome for MUT-NANOS1 (**Fig. 52A**).

Additionally, vinculin (VCL), another essential component that anchors AJs to the cytoskeleton via CTNNA1 and F-actin (Watabe-Uchida et al., 1998), was validated to be significantly downregulated (**Fig. 55B**), consistent with impaired cell-cell adhesion. Peng et al. (2010) demonstrated that VCL is vital for maintaining E-cadherin (CDH1) surface expression and AJ integrity. Their work showed that VCL depletion reduced E-cadherin-mediated adhesion, and this defect could be rescued by reintroducing VCL, but not by a VCL mutant incapable of binding β-catenin. This underscores the importance of the VCL–β-catenin interaction in stabilizing E-cadherin at the cell surface. Interestingly, both E-cadherin (CDH1) and VCL mRNAs are direct targets of MUT-NANOS1 binding, suggesting that this interaction may further repress their expression. Their significant downregulation likely results from a combinatorial effect of enhanced TGF-β signalling, possibly via TGFBR1-mediated canonical activation, and the additional repressive action of MUT-NANOS1.

Lastly, to further underscore the insufficient repression mediated by MUT-NANOS1, THBS1, an mRNA potentially bound by the protein and associated with the TGF-β pathway, was found to be upregulated (**Fig. 53** and **Fig. 55C–E**). THBS1 acts as a key mediator that reinforces TGF-β signalling through a positive feedback loop (Daubon et al., 2019).

In cancer studies, THBS1 acts as a primary physiological activator of TGF-β1, a potent inducer of EMT, promoting tumour invasion and migration (Massagué, 2012). Furthermore, THBS1 activates TGF-β1 by mobilizing its active form from the latent activating protein (LAP) via its type 1 domain (Schultz-Cherry et al., 1994). Interestingly, Daubon et al. (2019) found that in glioblastoma cells, treating with purified THBS1 did not increase fibronectin expression or induce Smad phosphorylation, and knocking down THBS1 did not alter TGF-β1 levels, suggesting that THBS1 may not universally activate TGF-β1. Further investigation by Daubon et al. (2019) revealed that THBS1 expression is upregulated by TGF-β1 through SMAD3 phosphorylation and nuclear translocation. They identified SMAD3 binding sites in the human THBS1 promoter, confirming that canonical TGF-β1 signalling via SMAD3 directly regulates THBS1 transcription. Thus, altered levels of THBS1 serve as reliable indicators of an activated or inactivated canonical TGF-β signalling pathway. Supporting this, Jayachandran et al. (2014) demonstrated that in melanoma cells, THBS1 promotes invasion likely through TGF-β activation. They showed that both THBS1 and TGF-β1 are expressed and secreted by mesenchymal-like melanoma cells, and that THBS1 is inducible in epithelial-like cells upon exposure to TGF-β1. This suggests a positive feedback loop where THBS1 activates latent TGF-β1, thereby promoting EMT and tumour progression. This could potentially serve to overcome the insufficient repression by MUT- NANOS1 due to the positive feedback loop mediated by activated TGF-ß/Smad signalling.

Although eCLIP data indicates that MUT-NANOS1 potentially binds to the 3'UTR of THBS1 mRNA, this interaction does not appear sufficient to supress THBS1 expression. Subsequent upregulation of the TGF-β pathway further drives THBS1 transcription via SMAD3-dependent mechanisms, thereby enhancing the availability of active TGF-β and promoting EMT and tumour progression as seen in cancer models (Daubon et al., 2019). In the current study, even if MUT-NANOS1 partially represses THBS1 mRNA, the feedback-driven elevation of THBS1 ensures continued TGF-β activation, effectively sustaining and amplifying TGF-β-driven processes despite the presence of a potentially repressive yet functionally compromised MUT-NANOS1 protein.

The coordinated dysregulation of THBS1, CDH1, and VCL in PGCs following MUT-NANOS1 overexpression parallels EMT-like processes seen in cancer cells, where elevated THBS1 and TGF-β activity prompt E-cadherin (CDH1) loss and cell-cell junction destabilization. This aberrant regulation underscores a common pathway through which MUT-NANOS1 influences cell behaviour, akin to mechanisms driving cancer progression. The mirrored EMT-like processes potentially disrupt the balance of PGC identity and cause cells to deviate toward a more somatic, possibly mesoderm-like lineage. The prolonged activation of TGF-β signalling, likely resulting from insufficient repression of targets such as TGFBR1 mRNA by MUT-NANOS1, may result in altered cell fate decisions and sustained canonical WNT signalling, as evidenced by the upregulation of certain components of the pathway such as LRP5 and BCL9L (**Fig. 54A-E**), which equally might have overridden the

dysfunctional repression mediated by binding of MUT-NANOS1. Compared to endogenous NANOS1, this insufficient repression underscores MUT-NANOS1's substantial impact on PGC differentiation *in vitro* and offers a possible mechanism by which it alters germ cell development.

MUT-NANOS1-mediated aberrant signalling correlates with reduced PGC markers and elevated ME markers Overexpression of MUT-NANOS1, and its associated aberrant repression of WNT, TGF-ß, and pluripotency-related markers leads to the downregulation of several critical PGC markers (**Fig. 56**), including SOX17, TFAP2C, NANOS3, OCT4 and NANOG. PRDM1 and PRDM14 are also significantly downregulated in MUT 3.2.5.2 cell line, although variability among biological replicates in the MUT 3.1.8 cell line precluded statistical significance for these markers (**Fig. 56**).

Importantly, SOX17 serves as the principal regulator of human PGC specification, orchestrating the expression of genes essential for germ cell identity (Irie et al., 2015). Similarly, PRDM1 supports PGCLC specification *in vitro* by repressing endodermal and other somatic gene markers, and it works in tandem with SOX17 to silence ME and somatic genes, thereby facilitating the establishment of PGC identity (Irie et al., 2015). Mutation-induced loss of PRDM1 impairs human PGC specification *in vitro*, even though SOX17 expression is maintained, highlighting PRDM1's crucial role downstream of SOX17 in maintaining germ cell integrity. Therefore, PRDM1, along with PRDM14 and TFAP2C, is a key determinant of PGC specification by repressing somatic genes (Ohinata et al., 2005; Magnúsdóttir et al., 2013).

Although critical for mouse PGC specification, PRDM14 shows delayed and significantly reduced expression in human PGCs *in vitro* and TCam-2 seminoma cells compared to hESCs (Grabole et al., 2013; Ma et al., 2011). Notably, PRDM14 is essential for maintaining pluripotency and facilitating the transition from pluripotent hESCs to differentiated PGCLCs, with its rapid downregulation and subsequent delayed re-expression being pivotal for PGC differentiation (Sybirna et al., 2020). PRDM14 depletion significantly reduces the efficiency of germ cell specification and results in an aberrant transcriptome in PRDM14-deficient human PGCs *in vitro*, resembling the effects observed upon loss of TFAP2C or PRDM1 (Sybirna et al., 2020).

TFAP2C, another essential marker for human PGC fate, is consistently and significantly downregulated in both MUT-NANOS1 cell lines (**Fig. 56**) and is crucial for maintaining germ cell identity as well as facilitating migration and gonadal colonization *in vivo* (Kojima et al., 2017; Sasaki et al., 2015; Li et al., 2017). The coordinated downregulation of TFAP2C and NANOS3 underscores the disruption of the germ cell specification pathway due to MUT-NANOS1 overexpression. TFAP2C collaborates with SOX17 and PRDM1 to regulate target genes essential for germ cell identity, including NANOS3, OCT4, and NANOG (Tang et al., 2022). Therefore, the downregulation of NANOS3 (**Fig. 56**), directly regulated by the cooperative action of SOX17, TFAP2C, and PRDM1, underscores this

disruption. As a downstream target, NANOS3 is essential for ensuring the survival and maintenance of germ cells by preventing apoptosis and supporting the germ cell program (Tang et al., 2022).

Notably, Tang et al. (2022) demonstrated that SOX17, TFAP2C, and PRDM1 collaboratively regulate a set of target genes essential for PGC identity by integrating signals and shaping the chromatin landscape towards human PGC fate. For instance, TFAP2C and SOX17 work together to upregulate and sustain the expression of OCT4 and NANOG, as well as transcriptional repressors like PRDM1 and CBFA2T2. TFAP2C promotes the expression of H3K9 demethylases that likely remove H3K9me2 marks and reorganize chromatin in human PGCs (Eguizabal et al., 2016). Additionally, in concert with PRDM1, it drives the expression of key components of the BAF chromatin remodeling complex, which is essential for maintaining lineage-specific enhancers (Alver et al., 2017).

In summary, SOX17 and TFAP2C initiate and sustain the expression of crucial transcription factors, including PRDM1, OCT4, and NANOG, which then cooperate with epigenetic remodelers to shape the chromatin landscape towards human PGC fate (Tang et al., 2022). Overexpressing MUT-NANOS1 disrupts this regulatory network undermining PGC identity, impairing the PGC program, and likely ultimately compromising PGCLC functionality.

On the other hand, in PGCLCs, the validated upregulation of ME markers was inconsistent between the two MUT-NANOS1 cell lines (**Fig. 60**). Only LEF1 and TBX3 were significantly upregulated in both cell lines, whereas other markers (e.g., WLS, HAND1, and BMP4) were significantly upregulated in only one cell line or not at all, suggesting a more specific cell line differentiation trajectory. Surprisingly, WNT5A was significantly downregulated, contrary to previous reports of its upregulation in late mesoderm lineages (Castillo-Venzor et al., 2023) and based on preliminary RNA-seq. DGE analysis in the MUT 3.2.5.2 cell line before and after MUT-NANOS1 overexpression in PGCs.

WNT5A is one of the several WNT ligands that play distinct roles in PGC development. Specifically, WNT3 and WNT3A are involved in PGC specification, corresponding to the PGCLCs examined in the current study (Aramaki et al., 2013; Tanaka et al., 2013), while WNT5A, together its receptor ROR2, is essential for PGC migration (Laird et al., 2011; Chawengsaksophak et al., 2012). The specification and sex differentiation of PGCs rely on the β-catenin–dependent canonical WNT signalling pathway (Aramaki et al., 2013), whereas PGC migration is regulated by the noncanonical WNT pathway mediated by WNT5A and ROR2 (Laird et al., 2011; Chawengsaksophak et al., 2012).

In cell lines, WNT5A has been shown to suppress canonical WNT signalling via ROR2 (Mikels and Nusse, 2006), an interaction confirmed *in vivo* by the observation that loss of ROR2 leads to increased canonical WNT activity (Mikels et al., 2009). Therefore, the aberrant upregulation of canonical WNT signalling due to overexpression of MUT-NANOS1 may have negatively affected the non-canonical WNT pathway potentially affecting expression of WNT5A (**Fig. 60**) and ROR2.

The WNT5A–ROR2 signalling axis also regulates PGC proliferation. In Ror2 mutant PGCs, elevated cell cycle gene expression and increased proliferation are observed, particularly in the hindgut where WNT5A expression is highest (Cantú et al., 2016). Molecular and genetic evidence suggests that WNT5A, acting via ROR2, suppresses β-catenin–dependent canonical WNT signalling in PGCs, thereby limiting their proliferation in specific locations. This balance between β-catenin–dependent and ROR2-mediated signalling fine-tunes the expansion and migration of PGCs, favouring early migrators while restricting latecomers (Cantú et al., 2016).

Given WNT5A's crucial role in facilitating PGC migration and regulating proliferation, its downregulation (**Fig. 60**) suggests that MUT-NANOS1 may additionally impair PGC migration by reducing WNT5A levels. This likely results from enhanced canonical WNT signalling activity, supported by upregulated TGF-ß signalling, which disrupts the delicate coordination between PGC movement and proliferation. Such an imbalance could lead to defects in germ cell numbers and subsequent migration through the hindgut toward the gonads, as demonstrated in infertility studies involving mutated WNT5A.

MUT-NANOS1 significantly reduces PGC numbers *in vitro* compared to WT-NANOS1 The significant downregulation of NANOS3 mRNA (**Fig. 56**) in both validated MUT-NANOS1 cell lines, where NANOS3 serves as the classic germ cell marker for FACS-based separation of PGCs from somatic cells, suggests a reduction in overall PGC numbers *in vitro* following inducible overexpression of MUT-NANOS1. Indeed, a statistically significant decrease was observed exclusively in the MUT cell line, whereas no significant change was noted for the WT counterpart (**Fig. 59**).

The MUT-NANOS1 p.[Pro34Thr; Ser78del] inadequate binding and potential GOF repression activity, leads to aberrant signalling that downregulates PGC and pluripotency markers, including NANOS3 mRNA, thereby significantly reducing the *in vitro* differentiation efficiency of human PGCs. This mechanistic disruption may help explain the *in vivo* phenotype, specifically the lack of gonadal PGCs in infertile male patients, who are dominant heterozygous carriers of the MUT-NANOS1 variant.

Inhibition of WNT signalling by Wnt-C59 inhibitor directs cell fate toward a PGC phenotype Further validation of the role of excessive WNT signalling in promoting a non-PGC phenotype was obtained by treating MUT-NANOS1–overexpressing PGCs with Wnt-C59, an inhibitor of Porcupine (PORCN). PORCN, a membrane-bound O-acyltransferase, is essential for WNT palmitoylation, secretion, and activity (Proffitt et al., 2013). Inhibiting PORCN blocks both canonical and noncanonical WNT signalling. Proffitt et al. (2013) showed that Wnt-C59 effectively inhibits PORCN activity at nanomolar concentrations by reducing WNT palmitoylation, impairing Wntless/WLS–mediated WNT secretion, and decreasing β-catenin reporter activity.

Consistent with these findings, WNT inhibition resulted in a significant surge in the levels of pluripotency markers OCT4 and NANOG, as well as NANOS3 and PRDM14 (**Fig. 61A**). PRDM14, an important marker toward the end of PGC specification (Sybirna et al., 2020), is also vital for hESC pluripotency (Chia et al., 2010). Notably, OCT4 and NANOG mRNAs potentially bound by MUT-NANOS1, were upregulated upon WNT inhibition, underscoring a reciprocal correlation between aberrant WNT signalling and the downregulation of pluripotency mediated by MUT-NANOS1. Furthermore, the upregulation of NANOS3 mRNA levels after WNT inhibition suggests an improvement in PGC induction efficiency *in vitro*, counteracting the negative impact observed with MUT-NANOS1 overexpression (**Fig. 59**).

Importantly, β-catenin, also likely bound by MUT-NANOS1, is significantly downregulated following WNT inhibition (**Fig. 61B**), suggesting that the previously upregulated canonical WNT pathway was effectively suppressed. Consequently, late mesoderm markers LEF1 and HAND1, which were notably upregulated MUT 3.2.5.2 cell line (**Fig. 60**), are significantly reduced following WNT inhibition (**Fig. 61B**), implying that cells progressing along the mesodermal differentiation trajectory, due to prolonged WNT signalling, may revert to a more PGC-like profile characterized by enhanced pluripotency.

In addition, TBX3, a downstream activator of WNT signalling (Price et al., 2012), also underwent significant downregulation following WNT inhibition (**Fig. 61B**), potentially further attenuating Nodal/SMAD2 signalling associated with TGF-ß pathway, given that Nodal is an intermediate target of TBX3 (Weidgang et al., 2013). This likely contributes to the suppression of the ME trajectory that PGCs seem to adopt under MUT-NANOS1 overexpression.

Lastly, BMP4 levels were significantly reduced following WNT inhibition (**Fig. 61B**), underscoring the interplay between BMP and WNT signalling, with WNT inhibition mediating the downregulation of BMP signalling.

It is important to note that the addition of the WNT inhibitor was synchronized with the induction of MUT-NANOS1 overexpression on Day 3 of the 4-day differentiation protocol. Consequently, the effect on early germ cell markers upstream of PRDM14, such as SOX17, PRDM1, and TFAP2C, may not have been temporally sufficient to induce statistically significant upregulation, although an increase in TFAP2C expression was observed (data not shown). Furthermore, BMP4, a well-studied signaling protein and member of the TGF-β family (also upregulated in MUT-NANOS1 PGCs), plays a crucial role in inducing mesoderm during gastrulation and in limiting PGC numbers after their specification.

In mouse embryonic development, BMP4 induces early nascent mesoderm and reaches peak expression in the extraembryonic mesoderm (ExM), amnion, and allantois, via its differentiated derivatives (Hadas et al., 2024). As chorion and PGC differentiation progress, BMP4 expression gradually decreases, mirroring the downregulation observed following WNT pathway inhibition (**Fig.**

61B). This reduction is significant because BMP4 signalling from the ExM at later stages effectively restricts the window of PGC specification by promoting differentiation toward the allantois in mice (Hadas et al., 2024). Moreover, Hadas et al. (2024), demonstrated that germline Bmp4 - knockout embryos exhibit reduced numbers of both allantois cells and PGCs, whereas ablating BMP4 exclusively from the embryo results in a marked increase in PGC frequency with a moderate yet significant reduction in allantois cells. These observations suggest that BMP4 not only induces mesoderm formation but also limits PGC numbers post-specification by directing cells toward mesodermal fates.

The interplay between BMP4 and WNT signalling is critical, as dynamic changes in BMP, WNT, and Nodal components are fundamental to germ-layer specification across mammals. Gastrulation is initiated by BMP signalling, which activates WNT signalling; in turn, WNT induces Nodal, and Nodal maintains BMP signalling, culminating in primitive streak formation (Chhabra et al., 2019). Therefore, the downregulation of BMP4 following WNT inhibition underscores the importance of the timing and interplay among these signals.

In primates, the amnion appears to substitute for the BMP4-producing extraembryonic ectoderm in mice, with BMP4 expression exhibiting similar temporal dynamics relative to WNT and Nodal expression (Yang et al., 2021; Hopf et al., 2011). Notably, significant downregulation of BMP4 signaling from the amnion in primates leads to a failure in mesoderm formation, as demonstrated by Yang et al. (2021) using a microfluidic-based embryonic sac model of amnion–epiblast interactions with ISL1-null hESCs.

In summary, BMP4 induces mesoderm formation and limits PGC numbers post-specification by favouring mesodermal derivatives. Its precise regulation during germ layer formation and subsequent reduction following WNT inhibition may reflect interrupted WNT-mediated Nodal signalling. Halting prolonged BMP signalling at the end of PGC specification helps restore normal germ cell marker expression. By blocking WNT signalling, Wnt-C59 partially reverses mesodermal differentiation, re-establishes PGC identity, and restores key pluripotency and germ cell markers. Simultaneous BMP4 downregulation disrupts the BMP–WNT–Nodal feedback loop, allowing cells to return to a PGC–specific pathway rather than progressing toward mesoderm.

SUMMARY OF RESULTS

Collectively, these findings indicate that MUT-NANOS1 p.[Pro34Thr; Ser78del] disrupts the pre-me RNA interactome via a premature, repressive GOF mechanism, potentially through direct motif recognition, altered protein interactions, thereby downregulating canonical WNT signaling (**Discussion Fig. 1**).

1. This reduction in WNT activity adversely affects key regulators such as EOMES, diminishing PGC competence in peri-gastrulation precursors. By disturbing the finely tuned signaling and transcriptional networks crucial for PGC specification, particularly through reduced WNT signaling, MUT-NANOS1 overexpression hinders PGC formation *in vitro*.

2. Furthermore, when overexpressed in differentiated PGCs, MUT-NANOS1 modulates essential signalling pathways (e.g. WNT, TGF-ß) and key transcription factors for PGC identity.

3. The concomitant upregulation of canonical WNT components (e.g., LRP5, BCL9L) and prolonged TGF-β signalling, coupled with insufficient repression of critical mRNAs by MUT-NANOS1, leads to aberrant activation of the canonical WNT pathway, EMT-like processes, and a shift toward mesendoderm and mesoderm fates.

4. This dysregulation is further compounded by the direct suppression of pluripotency factors OCT4 and NANOG, resulting in decreased expression of key germ cell markers (SOX17, PRDM1, PRDM14, TFAP2C, and NANOS3) and ultimately reducing the efficiency of PGC differentiation *in vitro*.

5. These alterations in gene expression and signalling dynamics provide a potential mechanistic explanation for the diminished number of PGCs observed *in vitro* and the absence of gonadal PGCs in infertile male patients harbouring the MUT-NANOS1 variant.

6. Notably, inhibiting WNT restores pluripotency and germ cell markers while reducing BMP4, underscoring the vital interplay between BMP, WNT, and TGF-β in maintaining proper germ cell fate.

7. Lastly, in this study, it appears that WT-NANOS1 does not play an essential role at this early stage of PGC specification, as its overexpression and minimal altered RNA interactome failed to significantly enhance PGC differentiation *in vitro*.

8. It is also highly probably that a knockdown of WT- NANOS1 would likely have little impact on differentiating PGCs at this point, suggesting that NANOS1's potential functions may become more relevant later in the differentiation trajectory. If more advanced *in vitro* germ cell models become available, it would be worthwhile to investigate whether NANOS1 has a significant role at later stages.

LIMITATIONS

There are certain technical and resource-based limitations to this study that should be considered when interpreting the effects of WT- and MUT-NANOS1 overexpression on hESC W15 cell line differentiation into pre-me and PGCLCs.

1.The enhanced crosslinking immunoprecipitation (eCLIP) experiments, aimed at delineating the RNA interactome for each NANOS1 protein counterpart, were conducted exclusively at the pre-me stage due to the high cell number requirement (approximately 20 million cells), which could not be met at the PGCLC stage. Although RNA-seq was performed for both stages, the bound interactome was inferred solely from pre-me, potentially underrepresenting targets that are either specific to or more prevalent in PGCLCs.

2. The PiggyBac-based cell lines employed for inducible overexpression of WT- and MUT-NANOS1 exhibited inherent differences in pluripotency. Specifically, the WT-NANOS1 cell line demonstrated reduced pluripotency relative to both the MUT-NANOS1 line and the unedited control (W15). This discrepancy, likely arising from technical challenges during single-cell sorting and clonal expansion, resulted in significant baseline differences in gene expression, thereby complicating the interpretation of WT-NANOS1's role in PGC specification.

3. While eCLIP is a robust method for mapping RNA-protein interactions, further validation, such as luciferase reporter assays, is required to conclusively confirm the direct binding and repressive effects of MUT-NANOS1 on its target mRNAs.

4. The PiggyBac overexpression system, doxycycline induction, and associated cell line generation protocols may have inadvertently influenced mRNA expression levels, as suggested by discrepancies observed in PCA analyses between the NANOS1-overexpressing cell lines and the control cell line W15.

5. Lastly, the differentiation model employed in this study represents only the PGC specification stage prior to the migration of PGCs across the hindgut to the gonads. It is possible that WT-NANOS1 does not play a major role at this early stage of PGC development, underscoring the need for more advanced models to fully elucidate the functions of both WT- and MUT-NANOS1 during later stages of germ cell development.

Collectively, these limitations highlight the necessity for further methodological refinements and the development of advanced differentiation models to comprehensively understand the mechanistic roles of endogenous NANOS1 and associated variants in germ cell specification and subsequent development.

REFERENCES

Albarqi, M. M. Y. & Ryder, S. P. The role of RNA-binding proteins in orchestrating germline development in Caenorhabditis elegans. *Front. Cell Dev. Biol.* **10**, 1094295 (2023).

Alves-Lopes, J. P. *et al.* Specification of human germ cell fate with enhanced progression capability supported by hindgut organoids. *Cell Rep.* **42**, 111907 (2023).

Arnold, S. J. & Robertson, E. J. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat. Rev. Mol. Cell Biol. 10, 91–103 (2009).

Arrizabalaga, G. & Lehmann, R. A Selective Screen Reveals Discrete Functional Domains in Drosophila Nanos. *Genetics* **153**, 1825–1838 (1999).

Arvola, R. M., Weidmann, C. A., Hall, T. M. T. & Goldstrohm, A. C. Combinatorial control of messenger RNAs by Pumilio, Nanos and Brain Tumor Proteins. *RNA Biol.* **14**, 1445–1456 (2017).

Asaoka-Taguchi, M., Yamada, M., Nakamura, A., Hanyu, K. & Kobayashi, S. Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos. *Nat. Cell Biol.* **1**, 431–437 (1999).

Baines, R. A. Neuronal homeostasis through translational control. *Mol. Neurobiol.* **32**, 113–121 (2005).

Barr, A. R. *et al.* DNA damage during S-phase mediates the proliferation-quiescence decision in the subsequent G1 via p21 expression. *Nat. Commun.* **8**, 14728 (2017).

Bhandari, D., Raisch, T., Weichenrieder, O., Jonas, S. & Izaurralde, E. Structural basis for the Nanos-mediated recruitment of the CCR4–NOT complex and translational repression. *Genes Dev.* **28**, 888–901 (2014).

Bonnomet, A. *et al.* The E-cadherin-repressed hNanos1 gene induces tumor cell invasion by upregulating MT1-MMP expression. *Oncogene* **27**, 3692–3699 (2008).

Campolo, F. et al. Essential Role of Sox2 for the Establishment and Maintenance of the Germ Cell Line. Stem Cells 31, 1408–1421 (2013).

Castillo-Venzor, A. *et al.* Origin and segregation of the human germline. *Life Sci. Alliance* **6**, e202201706 (2023).

Chen, D. *et al.* Germline competency of human embryonic stem cells depends on eomesodermin. *Biol. Reprod.* **97**, 850–861 (2017).
Chen, D. *et al.* Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. *Cell Rep.* **29**, 4568-4582.e5 (2019).

Choo, Y. & Klug, A. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. Proc. Natl. Acad. Sci. 91, 11168–11172 (1994).

Codino, A. et al. NANOS2 is a sequence-specific mRNA-binding protein that promotes transcript degradation in spermatogonial stem cells. iScience 24, 102762 (2021).

Collart, M. A. & Panasenko, O. O. The Ccr4–Not complex. *Gene* **492**, 42–53 (2012).

Cooke, H. J. & Saunders, P. T. K. Mouse models of male infertility. *Nat. Rev. Genet.* **3**, 790–801 (2002).

Cooper, T. A., Wan, L. & Dreyfuss, G. RNA and Disease. *Cell* **136**, 777–793 (2009).

Durcova-Hills, G., Tang, F., Doody, G., Tooze, R. & Surani, M. A. Reprogramming Primordial Germ Cells into Pluripotent Stem Cells. PLoS ONE 3, e3531 (2008).

Edwards, T. A., Pyle, S. E., Wharton, R. P. & Aggarwal, A. K. Structure of Pumilio Reveals Similarity between RNA and Peptide Binding Motifs. Cell 105, 281–289 (2001).

Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–6 (2013).

Gkountela, S. *et al.* The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. *Nat. Cell Biol.* **15**, 113–122 (2013).

Gerasimavicius, L., Livesey, B. J. & Marsh, J. A. Loss-of-function, gain-of-function and dominant-negative mutations have profoundly different effects on protein structure. *Nat. Commun.* **13**, 3895 (2022).

Ginter-Matuszewska, B. *et al.* NANOS1 and PUMILIO2 bind microRNA biogenesis factor GEMIN3, within chromatoid body in human germ cells. *Histochem. Cell Biol.* **136**, 279 (2011).

Hackett, J. A., Kobayashi, T., Dietmann, S. & Surani, M. A. Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. Stem Cell Rep. 8, 1645–1658 (2017).

Haraguchi, S. *et al.* nanos1: a mouse nanos gene expressed in the central nervous system is dispensable for normal development. *Mech. Dev.* **120**, 721–731 (2003).

Hassoun, R., Schwartz, P., Rath, D., Viebahn, C. & Männer, J. Germ layer differentiation during early hindgut and cloaca formation in rabbit and pig embryos. J. Anat. 217, 665–678 (2010).

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. Cell 146, 519–532 (2011).

Hayashi, K. et al. Offspring from Oocytes Derived from in Vitro Primordial Germ Cell–like Cells in Mice. Science 338, 971–975 (2012a).

Hayashi, Y., Saitou, M. & Yamanaka, S. Germline development from human pluripotent stem cells toward disease modeling of infertility. *Fertil. Steril.* **97**, 1250–1259 (2012b).

Hertig, A. T. *et al.* A thirteen-day human ovum studied histochemically. *Am. J. Obstet. Gynecol.* **76**, 1025–1043 (1958).

Ilaslan, E. *et al.* Distinct Roles of NANOS1 and NANOS3 in the Cell Cycle and NANOS3-PUM1-FOXM1 Axis to Control G2/M Phase in a Human Primordial Germ Cell Model. *Int. J. Mol. Sci.* **23**, 6592 (2022).

Irie, N. *et al.* SOX17 Is a Critical Specifier of Human Primordial Germ Cell Fate. *Cell* **160**, 253–268 (2015).

Irie, N., Sybirna, A. & Surani, M. A. What Can Stem Cell Models Tell Us About Human Germ Cell Biology? Curr. Top. Dev. Biol. 129, 25–65 (2018).

Janecki, D. M. *et al.* Human NANOS1 Represses Apoptosis by Downregulating Pro-Apoptotic Genes in the Male Germ Cell Line. *Int. J. Mol. Sci.* **21**, 3009 (2020).

Janic, A., Mendizabal, L., Llamazares, S., Rossell, D. & Gonzalez, C. Ectopic Expression of Germline Genes Drives Malignant Brain Tumor Growth in Drosophila. *Science* **330**, 1824–1827 (2010).

Jaruzelska, J. *et al.* Conservation of a Pumilio-Nanos complex from Drosophila germ plasm to human germ cells. *Dev. Genes Evol.* **213**, 120–126 (2003).

Jo, K. *et al.* Efficient differentiation of human primordial germ cells through geometric control reveals a key role for Nodal signaling. *eLife* **11**, e72811 (2022).

Joly, W., Chartier, A., Rojas-Rios, P., Busseau, I. & Simonelig, M. The CCR4 Deadenylase Acts with Nanos and Pumilio in the Fine-Tuning of Mei-P26 Expression to Promote Germline Stem Cell Self-Renewal. *Stem Cell Rep.* **1**, 411–424 (2013).

Julaton, V. T. A. & Pera, R. A. R. NANOS3 function in human germ cell development. *Hum. Mol. Genet.* **20**, 2238–2250 (2011).

Kadyrova, L. Y., Habara, Y., Lee, T. H. & Wharton, R. P. Translational control of maternal Cyclin B mRNA by Nanos in the Drosophila germline. *Development* **134**, 1519–1527 (2007).

Keuckelaere, E. D., Hulpiau, P., Saeys, Y., Berx, G. & Roy, F. van. Nanos genes and their role in development and beyond. *Cell. Mol. Life Sci.* **75**, 1929–1946 (2018).

Kobayashi, T. *et al.* Principles of early human development and germ cell program from conserved model systems. *Nature* **546**, 416–420 (2017).

Kobayashi, T. & Surani, M. A. On the origin of the human germline. Development **145**, dev150433 (2018).

Kojima, Y. *et al.* Evolutionarily Distinctive Transcriptional and Signaling Programs Drive Human Germ Cell Lineage Specification from Pluripotent Stem Cells. *Cell Stem Cell* **21**, 517-532.e5 (2017).

Kusz, K. M. *et al.* The highly conserved NANOS2 protein: testis-specific expression and significance for the human male reproduction. *Mol. Hum. Reprod.* **15**, 165–171 (2009a).

Kusz, K. *et al.* NANOS3 gene mutations in men with isolated sterility phenotype. *Mol. Reprod. Dev.* **76**, 804–804 (2009b).

Kusz-Zamelczyk, K. *et al.* Mutations of NANOS1, a human homologue of the Drosophila morphogen, are associated with a lack of germ cells in testes or severe oligo-astheno-teratozoospermia. *J. Méd. Genet.* **50**, 187 (2013).

Lai, F. & King, M. L. Repressive translational control in germ cells. *Mol. Reprod. Dev.* **80**, 665–676 (2013).

Lebedeva, L. A. *et al.* Transcriptional quiescence in primordial germ cells. *Crit. Rev. Biochem. Mol. Biol.* **53**, 579–595 (2018).

Lehmann, R. & Nüsslein-Volhard, C. The maternal gene nanos has a central role in posterior pattern formation of the Drosophila embryo. *Dev. (Camb., Engl.)* **112**, 679–91 (1991).

Lolicato, F. *et al.* Potential role of Nanos3 in maintaining the undifferentiated spermatogonia population. *Dev. Biol.* **313**, 725–738 (2008).

Magnúsdóttir, E. *et al.* A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nat. Cell Biol.* **15**, 905–915 (2013).

Meister, G. *et al.* Characterization of a nuclear 20S complex containing the survival of motor neurons (SMN) protein and a specific subset of spliceosomal Sm proteins. *Hum. Mol. Genet.* **9**, 1977–1986 (2000).

Mendjan, S. *et al.* NANOG and CDX2 Pattern Distinct Subtypes of Human Mesoderm during Exit from Pluripotency. *Cell Stem Cell* **15**, 310–325 (2014).

Meyts, E. R.-D., McGlynn, K. A., Okamoto, K., Jewett, M. A. S. & Bokemeyer, C. Testicular germ cell tumours. *Lancet* **387**, 1762–1774 (2016).

Miles, W. O. *et al.* Post-transcriptional gene expression control by NANOS is up-regulated and functionally important in pRb-deficient cells. *EMBO J.* **33**, 2201–2215 (2014).

Moon, K. R. *et al.* Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* **37**, 1482–1492 (2019).

Mourelatos, Z. *et al.* miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* **16**, 720–728 (2002).

Mulas, C., Kalkan, T. & Smith, A. NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from the Ground State. Stem Cell Rep. 9, 77–91 (2017).

Murata, Y. & Wharton, R. P. Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in drosophila embryos. *Cell* **80**, 747–756 (1995).

Nagirnaja, L., Aston, K. I. & Conrad, D. F. Genetic intersection of male infertility and cancer. *Fertil. Steril.* **109**, 20–26 (2018).

Nakamura, T. *et al.* A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* **537**, 57–62 (2016).

Ohinata, Y. et al. A Signaling Principle for the Specification of the Germ Cell Lineage in Mice. J. End--End Test. 137, 571–584 (2009).

Raisch, T. *et al.* Distinct modes of recruitment of the CCR4–NOT complex by Drosophila and vertebrate Nanos. *EMBO J.* **35**, 974–990 (2016).

Ramakrishna, N. B., Murison, K., Miska, E. A. & Leitch, H. G. Epigenetic Regulation during Primordial Germ Cell Development and Differentiation. *Sex. Dev.* **15**, 411–431 (2021).

Rosemann, J. *et al.* NANOS1 restricts oral cancer cell motility and TGF-ß signaling. *Eur. J. Cell Biol.* **103**, 151400 (2024).

Sada, A., Suzuki, A., Suzuki, H. & Saga, Y. The RNA-Binding Protein NANOS2 Is Required to Maintain Murine Spermatogonial Stem Cells. *Science* **325**, 1394–1398 (2009).

Saitou, M., Barton, S. C. & Surani, M. A. A molecular programme for the specification of germ cell fate in mice. *Nature* **418**, 293–300 (2002).

Saitou, M. & Hayashi, K. Mammalian in vitro gametogenesis. *Science* **374**, eaaz6830 (2021).

Sasaki, K. *et al.* Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. *Cell Stem Cell* **17**, 178–194 (2015).

Sasaki, K. *et al.* The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion. *Dev. Cell* **39**, 169–185 (2016).

Shaw, W. R., Armisen, J., Lehrbach, N. J. & Miska, E. A. The Conserved miR-51 microRNA Family Is Redundantly Required for Embryonic Development and Pharynx Attachment in Caenorhabditis elegans. *Genetics* **185**, 897–905 (2010).

Sivley, R. M., Dou, X., Meiler, J., Bush, W. S. & Capra, J. A. Comprehensive Analysis of Constraint on the Spatial Distribution of Missense Variants in Human Protein Structures. *Am. J. Hum. Genet.* **102**, 415–426 (2018).

Sonoda, J. & Wharton, R. P. Recruitment of Nanos to hunchback mRNA by Pumilio. *Genes Dev.* **13**, 2704–2712 (1999).

Strumane, K. *et al.* E-Cadherin Regulates Human Nanos1, which Interacts with p120ctn and Induces Tumor Cell Migration and Invasion. *Cancer Res.* **66**, 10007–10015 (2006).

Suzuki, A., Tsuda, M. & Saga, Y. Functional redundancy among Nanos proteins and a distinct role of Nanos2 during male germ cell development. *Development* **134**, 77–83 (2007).

Suzuki, A., Igarashi, K., Aisaki, K., Kanno, J. & Saga, Y. NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs. *Proc. Natl. Acad. Sci.* **107**, 3594–3599 (2010).

Suzuki, A., Saba, R., Miyoshi, K., Morita, Y. & Saga, Y. Interaction between NANOS2 and the CCR4-NOT Deadenylation Complex Is Essential for Male Germ Cell Development in Mouse. *PLoS ONE* **7**, e33558 (2012).

Tam, P. P. L. & Loebel, D. A. F. Gene function in mouse embryogenesis: get set for gastrulation. Nat. Rev. Genet. 8, 368–381 (2007).

Tang, W. W. C. *et al.* A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. *Cell* **161**, 1453–1467 (2015).

Tang, W. W. C., Kobayashi, T., Irie, N., Dietmann, S. & Surani, M. A. Specification and epigenetic programming of the human germ line. *Nat. Rev. Genet.* **17**, 585–600 (2016).

Tang, W. W. C. *et al.* Sequential enhancer state remodelling defines human germline competence and specification. *Nat. Cell Biol.* **24**, 448–460 (2022).

Teague, S., Yao, L. & Heemskerk, I. The many dimensions of germline competence. *Curr. Opin. Cell Biol.* **85**, 102259 (2023).

Teo, A. K. K. et al. Pluripotency factors regulate definitive endoderm specification through eomesodermin. Genes Dev. 25, 238–250 (2011).

Tsuda, M. *et al.* Conserved Role of nanos Proteins in Germ Cell Development. *Science* **301**, 1239–1241 (2003).

Tyser, R. C. V. *et al.* A spatially resolved single cell atlas of human gastrulation. *bioRxiv* 2020.07.21.213512 (2020) doi:10.1101/2020.07.21.213512.

Waddington, C. H. How Animals Develop. (2015) doi:10.4324/9781315665566.

Wahle, E. & Winkler, G. S. RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.* **1829**, 561–570 (2013).

Wang, C. & Lehmann, R. Nanos is the localized posterior determinant in Drosophila. Cell 66, 637–647 (1991).

Weidmann, C. A. et al. Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio. eLife 5, e17096 (2016).

Witschi, E. Migration of the germ cells of human embryos from the yolk sac to the primitive gonadal folds. *Contrib. Embryol.* **32**, 67–80 (1948).

Wu, X. *et al.* A NANOS3 mutation linked to protein degradation causes premature ovarian insufficiency. *Cell Death Dis.* **4**, e825 (2013).

Yokobayashi, S. *et al.* Clonal variation of human induced pluripotent stem cells for induction into the germ cell fate†. *Biol. Reprod.* **96**, 1154–1166 (2017).

Yoshida, M. et al. Conserved and divergent expression patterns of markers of axial development in eutherian mammals. Dev. Dyn. 245, 67–86 (2016).

Zheng, Y. *et al.* Controlled modelling of human epiblast and amnion development using stem cells. *Nature* **573**, 421–425 (2019).

Zhou, Z. et al. RNA Binding Protein Nanos2 Organizes Post-transcriptional Buffering System to Retain Primitive State of Mouse Spermatogonial Stem Cells. Dev. Cell **34**, 96–107 (2015).
Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

Bailey, T. L. *et al.* MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).

Chen, G. et al. Chemically defined conditions for human iPSC derivation and culture. Nat. Methods 8, 424–429 (2011).

Kobayashi, T. et al. Principles of early human development and germ cell program from conserved model systems. Nature 546, 416–420 (2017).

Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

Randolph, L. N., Bao, X., Zhou, C. & Lian, X. An all-in-one, Tet-On 3G inducible PiggyBac system for human pluripotent stem cells and derivatives. Sci. Rep. 7, 1549 (2017).

Park, M. A., Jung, H. S. & Slukvin, I. Genetic Engineering of Human Pluripotent Stem Cells Using PiggyBac Transposon System. *Curr. Protoc. Stem Cell Biol.* **47**, e63 (2018).

Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

Van Nostrand EL, et. al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods. 2016 Jun;13(6):508-14

Han, Z., Madhavan, B. K., Kaymak, S., Nawroth, P. & Kumar, V. A Fast and Reliable Method to Generate Pure, Single Cell-derived Clones of Mammalian Cells. Bio-Protoc. 12, (2022).

Wickham, H. ggplot2, Elegant Graphics for Data Analysis. *R* 189–201 (2016) doi:10.1007/978-3-319-24277-4_9.

Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov.* **2**, 100141 (2021).

Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* **16**, 284–287 (2012).

Yu, G., Wang, L.-G., Yan, G.-R. & He, Q.-Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608–609 (2014).

Adams, J. C. & Lawler, J. The thrombospondins. *Int. J. Biochem. Cell Biol.* **36**, 961–968 (2004).

Aksoy, I. *et al.* Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *EMBO J.* **32**, 938–953 (2013).

Alver, B. H. *et al.* The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nat. Commun.* **8**, 14648 (2017).

Alves-Lopes, J. P. et al. Specification of human germ cell fate with enhanced progression capability supported by hindgut organoids. Cell Rep. 42, 111907 (2023).

Ahmad, S. A. I., Anam, M. B., Istiaq, A., Ito, N. & Ohta, K. Tsukushi is essential for proper maintenance and terminal differentiation of mouse hippocampal neural stem cells. *Dev., Growth Differ.* **62**, 108–117 (2020).

Ahn, Y. *et al.* Multiple modes of Lrp4 function in modulation of Wnt/β-catenin signaling during tooth development. *Development* **144**, 2824–2836 (2017).

Ahn, Y., Sims, C., Logue, J. M., Weatherbee, S. D. & Krumlauf, R. Lrp4 and Wise interplay controls the formation and patterning of mammary and other skin appendage placodes by modulating Wnt signaling. *Development* **140**, 583–593 (2013).

Aramaki, S. *et al.* A Mesodermal Factor, T, Specifies Mouse Germ Cell Fate by Directly Activating Germline Determinants. *Dev. Cell* **27**, 516–529 (2013).

Arnold, S. J., Hofmann, U. K., Bikoff, E. K. & Robertson, E. J. Pivotal roles for eomesodermin during axis formation,epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* **135**, 501–511 (2008).

Ávila-González, D. *et al.* Unraveling the Spatiotemporal Human Pluripotency in Embryonic Development. *Front. Cell Dev. Biol.* **9**, 676998 (2021).

Bahrami, S. & Drabløs, F. Gene regulation in the immediate-early response process. Adv. Biol. Regul. 62, 37–49 (2016).

Beyer, T. A. *et al.* Switch Enhancers Interpret TGF-β and Hippo Signaling to Control Cell Fate in Human Embryonic Stem Cells. *Cell Rep.* **5**, 1611–1624 (2013).

Boroviak, T. *et al.* Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**, dev167833 (2018).

Cadigan, K. M. & Waterman, M. L. TCF/LEFs and Wnt Signaling in the Nucleus. *Cold Spring Harb. Perspect. Biol.* **4**, a007906 (2012).

Calder, A. *et al.* Lengthened G1 Phase Indicates Differentiation Status in Human Embryonic Stem Cells. *Stem Cells Dev.* **22**, 279–295 (2013).

Cantú, A. V., Altshuler-Keylin, S. & Laird, D. J. Discrete somatic niches coordinate proliferation and migration of primordial germ cells via Wnt signaling. *J. Cell Biol.* **214**, 215–229 (2016).

Cantú, A. V. & Laird, D. J. Primordial germ cell migration and the Wnt signaling pathway. *Anim. Reprod.* **14**, 89–101 (2017).

Carlson, C. B., Lawler, J. & Mosher, D. F. Thrombospondins: from structure to therapeutics: Structures of thrombospondins. *Cell. Mol. Life Sci.* **65**, 672–686 (2008).

Castillo-Venzor, A. *et al.* Origin and segregation of the human germline. *Life Sci. Alliance* **6**, e202201706 (2023).

Chen, D. et al. Germline competency of human embryonic stem cells depends on eomesodermin. Biol. Reprod. 97, 850–861 (2017).

Chen, D. et al. Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. Cell Rep. 29, 4568-4582.e5 (2019).

Cheng, H., Shang, D. & Zhou, R. Germline stem cells in human. *Signal Transduct. Target. Ther.* **7**, 345 (2022).

Chhabra, S., Liu, L., Goh, R., Kong, X. & Warmflash, A. Dissecting the dynamics of signaling events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in human gastruloids. *PLoS Biol.* **17**, e3000498 (2019).

Chia, N.-Y. *et al.* A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* **468**, 316–320 (2010).

Chu, L.-F. *et al.* Single-cell RNA-seq reveals novel regulators of human embryonic stem cell differentiation to definitive endoderm. *Genome Biol.* **17**, 173 (2016).

Collier, A. J. et al. Comprehensive Cell Surface Protein Profiling Identifies Specific Markers of Human Naive and Primed Pluripotent States. Cell Stem Cell 20, 874-890.e7 (2017).

Coronado, D. *et al.* A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res.* **10**, 118–131 (2013).

Dang, C. V. *et al.* The c-Myc target gene network. *Semin. Cancer Biol.* **16**, 253–264 (2006).

Daubon, T. *et al.* Deciphering the complex role of thrombospondin-1 in glioblastoma development. *Nat. Commun.* **10**, 1146 (2019).

Dominguez-Sola, D. *et al.* Non-transcriptional control of DNA replication by c-Myc. *Nature* **448**, 445–451 (2007).

Durruthy-Durruthy, J. *et al.* The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. *Nat. Genet.* **48**, 44–52 (2016).

Eckei, G., Böing, M., Brand-Saberi, B. & Morosan-Puopolo, G. Expression Pattern of Axin2 During Chicken Development. *PLoS ONE* **11**, e0163610 (2016).

Eguizabal, C. *et al.* Characterization of the Epigenetic Changes During Human Gonadal Primordial Germ Cells Reprogramming. *STEM CELLS* **34**, 2418–2428 (2016).

Enomoto, M. *et al.* Autonomous regulation of osteosarcoma cell invasiveness by Wnt5a/Ror2 signaling. *Oncogene* **28**, 3197–3208 (2009).

Faherty, N. *et al.* CCN2/CTGF increases expression of miR-302 microRNAs, which target the TGFβ type II receptor with implications for nephropathic cell phenotypes. *J. Cell Sci.* **125**, 5621–5629 (2012).

Fu, J., Jiang, M., Mirando, A. J., Yu, H.-M. I. & Hsu, W. Reciprocal regulation of Wnt and Gpr177/mouse Wntless is required for embryonic axis formation. *Proc. Natl. Acad. Sci.* **106**, 18598–18603 (2009).

Funa, N. S. et al. β-Catenin Regulates Primitive Streak Induction through Collaborative Interactions with SMAD2/SMAD3 and OCT4. Cell Stem Cell 16, 639–652 (2015).

Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–6 (2013).

Geng, S. *et al.* Balancing WNT signalling in early forebrain development: The role of LRP4 as a modulator of LRP6 function. *Front. Cell Dev. Biol.* **11**, 1173688 (2023).

Ginter-Matuszewska, B. *et al.* NANOS1 and PUMILIO2 bind microRNA biogenesis factor GEMIN3, within chromatoid body in human germ cells. *Histochem. Cell Biol.* **136**, 279 (2011).

Gong, Y. *et al.* LDL Receptor-Related Protein 5 (LRP5) Affects Bone Accrual and Eye Development. *Cell* **107**, 513–523 (2001).

Grabole, N. *et al.* Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep.* **14**, 629–637 (2013).

Guder, C. *et al.* The Wnt code: cnidarians signal the way. *Oncogene* **25**, 7450–7460 (2006).

Hadas, R. *et al.* Temporal BMP4 effects on mouse embryonic and extraembryonic development. *Nature* **634**, 652–661 (2024).

He, F. *et al.* Wnt5a regulates directional cell migration and cell proliferation via Ror2-mediated noncanonical pathway in mammalian palate development. *Development* **135**, 3871–3879 (2008).

Heemskerk, I. *et al.* Rapid changes in morphogen concentration control self-organized patterning in human embryonic stem cells. *eLife* **8**, e40526 (2019).

Gupta, S. *et al.* Emerging role of ZBTB7A as an oncogenic driver and transcriptional repressor. *Cancer Lett.* **483**, 22–34 (2020).

Henderson, J. K. *et al.* Preimplantation Human Embryos and Embryonic Stem Cells Show Comparable Expression of Stage-Specific Embryonic Antigens. *STEM CELLS* **20**, 329–337 (2002).

Hopf, C., Viebahn, C. & Püschel, B. BMP signals and the transcriptional repressor BLIMP1 during germline segregation in the mammalian embryo. *Dev. Genes Evol.* **221**, 209 (2011).

Hou, J., Wen, Y.-H., Feng, K.-N., Ma, X.-F. & Yao, J.-P. DACT1 is involved in human placenta development by promoting Wnt signaling. *Arch. Gynecol. Obstet.* **291**, 1289–1296 (2015).

Irie, N. et al. SOX17 Is a Critical Specifier of Human Primordial Germ Cell Fate. Cell 160, 253–268 (2015).

Istiaq, A. & Ohta, K. A review on Tsukushi: mammalian development, disorders, and therapy. *J. Cell Commun. Signal.* **16**, 505–513 (2022).

Hadjihannas, M. V., Bernkopf, D. B., Brückner, M. & Behrens, J. Cell cycle control of Wnt/β-catenin signalling by conductin/axin2 through CDC20. *EMBO Rep.* **13**, 347–354 (2012).

Hart, A. H., Hartley, L., Ibrahim, M. & Robb, L. Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev. Dyn.* **230**, 187–198 (2004).

Hatano, S. *et al.* Pluripotential competence of cells associated with Nanog activity. *Mech. Dev.* **122**, 67–79 (2005).

Hernández-Martínez, R., Ramkumar, N. & Anderson, K. V. p120-catenin regulates WNT signaling and EMT in the mouse embryo. Proc. Natl. Acad. Sci. 116, 16872–16881 (2019).

Hernández-Martínez, R. *et al.* Axin1 and Axin2 regulate the WNT-signaling landscape to promote distinct mesoderm programs. *bioRxiv* 2024.09.11.612342 (2024) doi:10.1101/2024.09.11.612342.

Huelsken, J. *et al.* Requirement for β-Catenin in Anterior-Posterior Axis Formation in Mice. *J. Cell Biol.* **148**, 567–578 (2000).

Hou, J., Wen, Y.-H., Feng, K.-N., Ma, X.-F. & Yao, J.-P. DACT1 is involved in human placenta development by promoting Wnt signaling. *Arch. Gynecol. Obstet.* **291**, 1289–1296 (2015).

Huber, A. H., Stewart, D. B., Laurents, D. V., Nelson, W. J. & Weis, W. I. The Cadherin Cytoplasmic Domain Is Unstructured in the Absence of β-Catenin. *J. Biol. Chem.* **276**, 12301–12309 (2001).

Hudson, B. I. & Lippman, M. E. Targeting RAGE Signaling in Inflammatory Disease. *Annu. Rev. Med.* **69**, 1–16 (2017).

Ikeda, S. *et al.* Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. *EMBO J.* **17**, 1371–84 (1998).

Irie, N., Tang, W. W. C. & Surani, M. A. Germ cell specification and pluripotency in mammals: a perspective from early embryogenesis. *Reprod. Med. Biol.* **13**, 203–215 (2014).

Jayachandran, A. *et al.* Thrombospondin 1 promotes an aggressive phenotype through epithelial-to-mesenchymal transition in human melanoma. *Oncotarget* **5**, 5782–5797 (2014).

Jo, K. *et al.* Efficient differentiation of human primordial germ cells through geometric control reveals a key role for Nodal signaling. *eLife* **11**, e72811 (2022).

Johnson, E. B., Hammer, R. E. & Herz, J. Abnormal development of the apical ectodermal ridge and polysyndactyly in Megf7-deficient mice. *Hum. Mol. Genet.* **14**, 3523–3538 (2005).

Jostes, S. V. *et al.* Unique and redundant roles of SOX2 and SOX17 in regulating the germ cell tumor fate. *Int. J. Cancer* **146**, 1592–1605 (2020).

Kehler, J. *et al.* Oct4 is required for primordial germ cell survival. *EMBO Rep.* **5**, 1078–1083 (2004).

Kelly, O. G., Pinson, K. I. & Skarnes, W. C. The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. *Development* **131**, 2803–2815 (2004).

Kim, H. *et al.* Modulation of β-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. *Nat. Commun.* **4**, 2403 (2013).

Kobayashi, T. *et al.* Principles of early human development and germ cell program from conserved model systems. *Nature* **546**, 416–420 (2017).

Kojima, Y. *et al.* Evolutionarily Distinctive Transcriptional and Signaling Programs Drive Human Germ Cell Lineage Specification from Pluripotent Stem Cells. *Cell Stem Cell* **21**, 517-532.e5 (2017).

Kusz-Zamelczyk, K. et al. Mutations of NANOS1, a human homologue of the Drosophila morphogen, are associated with a lack of germ cells in testes or severe oligo-astheno-teratozoospermia. J. Méd. Genet. 50, 187 (2013).

Julaton, V. T. A. & Pera, R. A. R. NANOS3 function in human germ cell development. *Hum. Mol. Genet.* **20**, 2238–2250 (2011).

Laird, D. J., Altshuler-Keylin, S., Kissner, M. D., Zhou, X. & Anderson, K. V. Ror2 Enhances Polarity and Directional Migration of Primordial Germ Cells. *PLoS Genet.* **7**, e1002428 (2011).

Lecarpentier, Y., Schussler, O., Hébert, J.-L. & Vallée, A. Multiple Targets of the Canonical WNT/β-Catenin Signaling in Cancers. *Front. Oncol.* **9**, 1248 (2019).

Lee, Y., Swencki, B., Shoichet, S. & Shivdasani, R. A. A Possible Role for the High Mobility Group Box Transcription Factor Tcf-4 in Vertebrate Gut Epithelial Cell Differentiation*. *J. Biol. Chem.* **274**, 1566–1572 (1999).

Leung, J. Y. *et al.* Activation of AXIN2 Expression by β-Catenin-T Cell Factor A FEEDBACK REPRESSOR PATHWAY REGULATING Wnt SIGNALING*. *J. Biol. Chem.* **277**, 21657–21665 (2002).

Li, L. *et al.* Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell* **20**, 858-873.e4 (2017).

Liao, S. *et al.* Relationship between SDC1 and cadherin signalling activation in cancer. *Pathol. - Res. Pr.* **216**, 152756 (2020).

Lin, H. *et al.* sFRP2 activates Wnt/β-catenin signaling in cardiac fibroblasts: differential roles in cell growth, energy metabolism, and extracellular matrix remodeling. *Am. J. Physiol.-Cell Physiol.* **311**, C710–C719 (2016).

Kobayashi, T., Alberio, R. & Surani, M. A. Simulating gastrulation development and germ cell fate in vitro using human and monkey pluripotent stem cells. Protoc. Exch. (2017) doi:10.1038/protex.2017.050.

Kuehn, M. R. *et al.* Nodal Signaling Uses Activin and Transforming Growth Factor-β Receptor-regulated Smads*. *J. Biol. Chem.* **276**, 656–661 (2001).

Li, Y. *et al.* LRP4 Mutations Alter Wnt/β-Catenin Signaling and Cause Limb and Kidney Malformations in Cenani-Lenz Syndrome. *Am. J. Hum. Genet.* **86**, 696–706 (2010).

Lin, J., Zhang, Q., Zhu, L. Q., Yu, Q. H. & Yang, Q. The copy number and integration site analysis of IGF-1 transgenic goat. Int. J. Mol. Med. 34, 900–910 (2014).

Liu, J. *et al.* Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal Transduct. Target. Ther.* **7**, 3 (2022).

Looijenga, L. H. J., Stoop, H. & Biermann, K. Testicular cancer: biology and biomarkers. *Virchows Arch.* **464**, 301–313 (2014).

Lu, R., Yang, A. & Jin, Y. Dual Functions of T-Box 3 (Tbx3) in the Control of Self-renewal and Extraembryonic Endoderm Differentiation in Mouse Embryonic Stem Cells*. *J. Biol. Chem.* **286**, 8425–8436 (2011).

Ma, J. *et al.* miR-140-y targets TCF4 to regulate the Wnt signaling pathway and promote embryonic feather follicle development in Hungarian white goose. *Poult. Sci.* **103**, 103508 (2024).

Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A. & Wysocka, J. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nat. Struct. Mol. Biol.* **18**, 120–127 (2011).

Magnúsdóttir, E. *et al.* A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nat. Cell Biol.* **15**, 905–915 (2013).

Martinato, F., Cesaroni, M., Amati, B. & Guccione, E. Analysis of Myc-Induced Histone Modifications on Target Chromatin. *PLoS ONE* **3**, e3650 (2008).

Massagué, J. TGFβ signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**, 616–630 (2012).

Mikels, A. J. & Nusse, R. Purified Wnt5a Protein Activates or Inhibits β-Catenin–TCF Signaling Depending on Receptor Context. *PLoS Biol.* **4**, e115 (2006).

Mikels, A., Minami, Y. & Nusse, R. Ror2 Receptor Requires Tyrosine Kinase Activity to Mediate Wnt5A Signaling*. *J. Biol. Chem.* **284**, 30167–30176 (2009).

Mohammed, M. K. *et al.* Wnt/β-catenin signaling plays an ever-expanding role in stem cell self-renewal, tumorigenesis and cancer chemoresistance. *Genes Dis.* **3**, 11–40 (2016).

Molenaar, M. *et al.* Differential expression of the HMG box transcription factors XTcf-3 and XLef-1 during early Xenopus development. *Mech. Dev.* **75**, 151–154 (1998).

Niehrs, C. & Shen, J. Regulation of Lrp6 phosphorylation. *Cell. Mol. Life Sci.* **67**, 2551–2562 (2010).

Niimori, D. *et al.* Tsukushi controls the hair cycle by regulating TGF-β1 signaling. *Dev. Biol.* **372**, 81–87 (2012).

Nishioka, N. *et al.* The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. *Dev. Cell* **16**, 398–410 (2009).

Nishita, M. *et al.* Filopodia formation mediated by receptor tyrosine kinase Ror2 is required for Wnt5a-induced cell migration. *J. Cell Biol.* **175**, 555–562 (2006).

Niwa, H., Ogawa, K., Shimosato, D. & Adachi, K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460**, 118–122 (2009).

Oda, H. & Takeichi, M. Structural and functional diversity of cadherin at the adherens junction. J. Cell Biol. 193, 1137–1146 (2011).

Park, M. A., Jung, H. S. & Slukvin, I. Genetic Engineering of Human Pluripotent Stem Cells Using PiggyBac Transposon System. *Curr. Protoc. Stem Cell Biol.* **47**, e63 (2018).

Polyak, K. & Weinberg, R. A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat. Rev. Cancer* **9**, 265–273 (2009).

Price, F. D. *et al.* Canonical Wnt Signaling Induces a Primitive Endoderm Metastable State in Mouse Embryonic Stem Cells. *STEM CELLS* **31**, 752–764 (2013).

Proffitt, K. D. *et al.* Pharmacological Inhibition of the Wnt Acyltransferase PORCN Prevents Growth of WNT-Driven Mammary Cancer. *Cancer Res.* **73**, 502–507 (2013).

Prühs, R., Beermann, A. & Schröder, R. The Roles of the Wnt-Antagonists Axin and Lrp4 during Embryogenesis of the Red Flour Beetle Tribolium castaneum. *J. Dev. Biol.* **5**, 10 (2017).

Ohinata, Y. *et al.* Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**, 207–213 (2005).

Ohta, K. *et al.* Tsukushi functions as a Wnt signaling inhibitor by competing with Wnt2b for binding to transmembrane protein Frizzled4. *Proc. Natl. Acad. Sci.* **108**, 14962–14967 (2011).

Okamura, D., Tokitake, Y., Niwa, H. & Matsui, Y. Requirement of Oct3/4 function for germ cell specification. *Dev. Biol.* **317**, 576–584 (2008).

Richter, J. D. & Sonenberg, N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* **433**, 477–480 (2005).

Ring, A., Kim, Y.-M. & Kahn, M. Wnt/Catenin Signaling in Adult Stem Cell Physiology and Disease. *Stem Cell Rev. Rep.* **10**, 512–525 (2014).

Pauklin, S. & Vallier, L. The Cell-Cycle State of Stem Cells Determines Cell Fate Propensity. *Cell* **155**, 135–147 (2013).

Peng, X., Cuff, L. E., Lawton, C. D. & DeMali, K. A. Vinculin regulates cell-surface E-cadherin expression by binding to β-catenin. *J. Cell Sci.* **123**, 567–577 (2010).

Rao, J. *et al.* Stepwise Clearance of Repressive Roadblocks Drives Cardiac Induction in Human ESCs. *Cell Stem Cell* **18**, 554–556 (2016).

Rennoll, S. A., Scott, S. A. & Yochum, G. S. Targeted repression of AXIN2 and MYC gene expression using designer TALEs. *Biochem. Biophys. Res. Commun.* **446**, 1120–1125 (2014).

Reyes, M., Flores, T., Betancur, D., Peña-Oyarzún, D. & Torres, V. A. Wnt/β-Catenin Signaling in Oral Carcinogenesis. *Int. J. Mol. Sci.* **21**, 4682 (2020).

Roël, G., Gent, Y. Y. J., Peterson-Maduro, J., Verbeek, F. J. & Destree, O. Lef1 plays a role in patterning the mesoderm and ectoderm in Xenopus tropicalis. *Int. J. Dev. Biol.* **53**, 81–89 (2009).

Rosemann, J. et al. NANOS1 restricts oral cancer cell motility and TGF-ß signaling. Eur. J. Cell Biol. 103, 151400 (2024).

Schäfer, S. *et al.* The role of BLIMP1 and its putative downstream target TFAP2C in germ cell development and germ cell tumours. *Int. J. Androl.* **34**, e152–e159 (2011).

Sannino, G. *et al.* Role of BCL9L in transforming growth factor-β (TGF-β)-induced epithelial-to-mesenchymal-transition (EMT) and metastasis of pancreatic cancer. *Oncotarget* **7**, 73725–73738 (2016).

Salinas, P. C. & Zou, Y. Wnt Signaling in Neural Circuit Assembly. *Annu. Rev. Neurosci.* **31**, 339–358 (2008).

Sasaki, K. *et al.* Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. *Cell Stem Cell* **17**, 178–194 (2015).

Schultz-Cherry, S., Lawler, J. & Murphy-Ullrich, J. E. The type 1 repeats of thrombospondin 1 activate latent transforming growth factor-beta. *J. Biol. Chem.* **269**, 26783–8 (1994).

Slusarski, D. C., Yang-Snyder, J., Busa, W. B. & Moon, R. T. Modulation of embryonic intracellular Ca2+ signaling by Wnt-5A. *Dev. Biol.* **182**, 114–20 (1997).

Sekiya, T. *et al.* Identification of BMP and Activin Membrane-bound Inhibitor (BAMBI), an Inhibitor of Transforming Growth Factor-β Signaling, as a Target of the β-Catenin Pathway in Colorectal Tumor Cells*. *J. Biol. Chem.* **279**, 6840–6846 (2004).

Shtutman, M. *et al.* The cyclin D1 gene is a target of the β-catenin/LEF-1 pathway. *Proc. Natl. Acad. Sci.* **96**, 5522–5527 (1999).

Sierra, R. A. et al. TCF7L1 suppresses primitive streak gene expression to support human embryonic stem cell pluripotency. Development 145, dev161075 (2018).

Sonoda, J. & Wharton, R. P. Recruitment of Nanos to hunchback mRNA by Pumilio. *Genes Dev.* **13**, 2704–2712 (1999).

Su, Z. *et al.* Antagonism between the transcription factors NANOG and OTX2 specifies rostral or caudal cell fate during neural patterning transition. *J. Biol. Chem.* **293**, 4445–4455 (2018).

Sybirna, A. *et al.* A critical role of PRDM14 in human primordial germ cell fate revealed by inducible degrons. *Nat. Commun.* **11**, 1282 (2020).

Rahl, P. B. *et al.* c-Myc Regulates Transcriptional Pause Release. *Cell* **141**, 432–445 (2010).

Riggelen, J. van, Yetil, A. & Felsher, D. W. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat. Rev. Cancer* **10**, 301–309 (2010).

Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).

Tanaka, S. S. *et al.* Dullard/Ctdnep1 Modulates WNT Signalling Activity for the Formation of Primordial Germ Cells in the Mouse Embryo. *PLoS ONE* **8**, e57428 (2013).

Teague, S., Yao, L. & Heemskerk, I. The many dimensions of germline competence. *Curr. Opin. Cell Biol.* **85**, 102259 (2023).

Terasaki, H., Saitoh, T., Shiokawa, K. & Katoh, M. Frizzled-10, up-regulated in primary colorectal cancer, is a positive regulator of the WNT - beta-catenin - TCF signaling pathway. Int. J. Mol. Med. 9, 107–12 (2002).

Tang, W. W. C. et al. Sequential enhancer state remodelling defines human germline competence and specification. Nat. Cell Biol. 24, 448–460 (2022).

Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139**, 871–890 (2009).

Tyser, R. C. V. *et al.* Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* **600**, 285–289 (2021).

Wal, T. van der & Amerongen, R. van. Walking the tight wire between cell adhesion and WNT signalling: a balancing act for β-catenin. Open Biol. 10, 200267 (2020).

Vallier, L. *et al.* Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* **136**, 1339–1349 (2009).

Wang, Q., Yang, L., Alexander, C. & Temple, S. The Niche Factor Syndecan-1 Regulates the Maintenance and Proliferation of Neural Progenitor Cells during Mammalian Cortical Development. *PLoS ONE* **7**, e42883 (2012).

Watabe-Uchida, M. *et al.* α-Catenin-Vinculin Interaction Functions to Organize the Apical Junctional Complex in Epithelial Cells. *J. Cell Biol.* **142**, 847–857 (1998).

Weidgang, C. E. *et al.* TBX3 Directs Cell-Fate Decision toward Mesendoderm. *Stem Cell Rep.* **2**, 747 (2014).

Wey, A. & Knoepfler, P. S. c-myc and N-myc promote active stem cell metabolism and cycling as architects of the developing brain. *Oncotarget* **1**, 120–130 (2010).

Willnow, T. E., Christ, A. & Hammes, A. Endocytic receptor-mediated control of morphogen signaling. *Development* **139**, 4311–4319 (2012).

Wright, M. H. et al. Brca1 breast tumors contain distinct CD44+/CD24- and CD133+cells with cancer stem cell characteristics. Breast Cancer Res. 10, R10 (2008).

Xu, R.-H. *et al.* NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell stem cell* **3**, 196–206 (2007).

Yagi, R. *et al.* Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* **134**, 3827–3836 (2007).

Yang, R. *et al.* Amnion signals are essential for mesoderm formation in primates. *Nat. Commun.* **12**, 5126 (2021).

Yap, A. S., Brieher, W. M. & Gumbiner, B. M. Molecular and functional analysis of cadherin-based adherens junctions. *Annu. Rev. cell Dev. Biol.* **13**, 119–146 (1997).

Yeom, Y. I. *et al.* Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881–894 (1996).

Yochum, G. S., Cleland, R. & Goodman, R. H. A Genome-Wide Screen for β-Catenin Binding Sites Identifies a Downstream Enhancer Element That Controls c-Myc Gene Expression. *Mol. Cell. Biol.* **28**, 7368–7379 (2008).

Yu, S. *et al.* BMP4 resets mouse epiblast stem cells to naive pluripotency through ZBTB7A/B-mediated chromatin remodelling. *Nat. Cell Biol.* **22**, 651–662 (2020).

Zhang, P., Andrianakos, R., Yang, Y., Liu, C. & Lu, W. Kruppel-like Factor 4 (Klf4) Prevents Embryonic Stem (ES) Cell Differentiation by Regulating Nanog Gene Expression. *J. Biol. Chem.* **285**, 9180–9189 (2010).

Zhang, Y., Sivasankar, S., Nelson, W. J. & Chu, S. Resolving cadherin interactions and binding cooperativity at the single-molecule level. Proc. Natl. Acad. Sci. 106, 109–114 (2009).

Zhijun, Z. & Jingkang, H. MicroRNA-520e suppresses non-small-cell lung cancer cell growth by targeting Zbtb7a-mediated Wnt signaling pathway. *Biochem. Biophys. Res. Commun.* **486**, 49–56 (2017).

Zhou, X., Vink, M., Klaver, B., Berkhout, B. & Das, A. T. Optimization of the Tet-On system for regulated gene expression through viral evolution. Gene Ther. 13, 1382–1390 (2006).