



Identification of transcriptional modulators of noise and differentiation in embryonic stem cells

Dissertation

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

der Fakultät für Chemie und Chemische Biologie

der Technischen Universität Dortmund

Angefertigt am Max-Planck-Institut für molekulare Physiologie in Dortmund

vorgelegt von

Max Fernkorn

Februar 2024

Date of submission: 26.02.2024

1st Examiner: Prof. Dr. Philippe Bastiaens

2nd Examiner: Prof. Dr. Jan Hengstler

Supervisor: Dr. Christian Schröter

The work described in this thesis was performed under the supervision of Dr. Christian Schröter and the formal supervision of Prof. Dr. Philippe Bastiaens at the Department of Systemic Cell Biology, Max Planck Institute of Molecular Physiology, Dortmund, Germany.

Max Fernkorn was affiliated with the International Max Planck Research School for Living Matter, Dortmund, Germany.

The work described in this thesis has been partly included in the following publication:

Fernkorn M., Schröter C. 2024. *Med12* cooperates with multiple differentiation signals to enhance embryonic stem cell plasticity. bioRxiv doi.org/10.1101/2024.01.22.576603

The author also contributed to the following publication during the Ph.D. studies:

Schumacher S., Fernkorn M., Marten M., Kim Y. S., Bedzhov I., Schröter C. 2023. Tissue-intrinsic Wnt signals antagonize Nodal-driven AVE differentiation. bioRxiv doi.org/10.1101/2023.05.19.54143

Declaration

Eidesstattliche Versicherung (Affidavit)

Fernkorn, Max

Name, Vorname
(Surname, first name)

216863

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Dortmund, 26.02.2024

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Identification of transcriptional modulators of noise and differentiation

in embryonic stem cells

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.

Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Dortmund, 26.02.2024

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Abstract

During embryonic development, many cell types arise from a single homogenous cell population. Coordination via cell-cell signaling leads to functional specialization by changing the set of expressed genes. In very early mammalian development, the homogenous cell population of pluripotent cells in the inner cell mass breaks symmetry and differentiates into epiblast and primitive endoderm lineages under the influence of Fibroblast Growth Factor (FGF), Wnt, and mammalian Target Of Rapamycin (mTOR) signaling. These signals modulate the activity of sequence-specific and general transcription factors. The involved sequence-specific transcription factors have not been identified for all mentioned signals and it is unclear whether there are specific components of the general transcriptional machinery which are especially important during early differentiation. Moreover, it is an open question how signals and general transcription factors cooperate to shape the transcriptional dynamics.

Here, I ask how developmental signals are translated into transcriptional responses in pluripotent mouse embryonic stem cells (mESCs). I show that FGF4 signaling affects the transcriptional dynamics at target genes, resulting in an increase in cellular noise. I perform a genome-wide CRISPR screen to identify further signaling components and general transcriptional regulators for gene expression changes upon differentiation signals. Specifically, I focus on gene perturbations affecting the expression of a *Spry4*^{H2B-Venus} reporter, which is upregulated in response to FGF signaling during the transition from naïve pluripotency to both epiblast and primitive endoderm cells. This screen returns multiple hits related to FGF signaling as expected and reveals general transcription factors associated with the Elongator and Mediator complexes. Focusing on *Med12*, a Mediator subunit whose loss affects the expression of the *Spry4* reporter strongest, I find that it regulates gene expression during pluripotency transitions acting in parallel to FGF, Wnt, and mTOR signaling. During the exit of pluripotency, *Med12*-mutant cells react less efficiently to changes in the signaling environment both functionally and on the gene expression level. Surprisingly, the generation of epiblast and primitive endoderm cells is largely buffered against the loss of *Med12*. During this bifurcation, *Med12*-mutant cells show lower plasticity and noise levels, causing a better separation between the two cell types.

These findings suggest that *Med12* is an important general transcription factor during differentiation to prime transcriptional changes. *Med12* acts in parallel to FGF-signaling in order to regulate the transcriptional variability and thereby allows cells to explore differentiation trajectories while keeping their plasticity.

Zusammenfassung

Während der Embryonalentwicklung entstehen verschiedene Zelltypen aus einer einzigen homogenen Zellpopulation. Welche Gene für die funktionale Spezialisierung exprimiert werden, wird durch Zell-Zell Kommunikation koordiniert. In sehr frühen Stadien der Embryonalentwicklung von Säugetieren tragen die Signalmoleküle FGF (*Fibroblast Growth Factor*), Wnt (*Wingless Integrated*) und mTOR (*mammalian Target Of Rapamycin*) dazu bei, Zellen der inneren Zellmasse zu entweder Epiblast- oder primitiven Endodermzellen zu differenzieren. Diese Signale beeinflussen die Aktivität von sequenz-spezifischen und generellen Transkriptionsfaktoren. Nicht für alle involvierten Signale wurden bereits sequenz-spezifische Faktoren identifiziert. Zudem ist unbekannt welche generellen Komponenten der Transkriptionsmaschinerie besonders wichtig während der frühen Zelldifferenzierung sind und wie diese mit den Signalen kooperieren um die Dynamik der Transkription zu regulieren.

In dieser Arbeit erforsche ich anhand von embryonalen Mäusestammzellen wie die Signale während der Embryonalentwicklung in transkriptionelle Antworten übersetzt werden. Die Ergebnisse zeigen, dass FGF4 die Dynamik der Transkription so reguliert, dass die Zielgene eine besonders hohe Variabilität aufweisen. Durch ein genomweites CRISPR Mutantenscreening identifiziere ich verschiedene Signalwege und allgemeine transkriptionelle Regulatoren, die die Genexpression verändern. Speziell konzentriere ich mich auf Mutationen, die ein fluoreszentes *Spry4*-Reportergen regulieren, das während des Übergangs von naiver Pluripotenz zu Epiblast- und primitiven Endodermzellen exprimiert wird. Dieses Screening bestätigt den Zusammenhang mit dem FGF-Signalweg und offenbart allgemeine Transkriptionsfaktoren, die mit den Elongator- und Mediator-Komplexen assoziiert sind. Den größten Effekt auf die Expression des *Spry4*-Reportergens hat der Verlust von *Med12*. Diese Untereinheit des Mediator-Komplexes beeinflusst die Genexpression beim Übergang zwischen verschiedenen Pluripotenzstadien parallel zu FGF-, Wnt- und mTOR-Signalen. *Med12*-mutante Stammzellen reagieren sowohl funktionell als auch auf der Ebene der Genexpression weniger effizient auf Veränderungen in der Signalumgebung. Überraschenderweise bleibt die Generierung von Epiblast- und primitiven Endodermzellen weitgehend unverändert zwischen Wildtyp und *Med12*-mutanten Zellen. Während dieser Verzweigung zeigen *Med12*-mutierte Zellen niedrigere Plastizitäts- und Variabilitätsniveaus, was zu einer besseren Trennung zwischen den beiden Zelltypen führt.

Die Ergebnisse legen nahe, dass *Med12* ein wichtiger genereller Transkriptionsfaktor während der Zelldifferenzierung ist um transkriptionelle Veränderungen zu initiieren. Zusammen mit dem parallel agierendem FGF-Signal ermöglicht *Med12* Zellen die Variabilität der Transkription zu regulieren, sodass diese verschiedene Differenzierungspfade erkunden können während ihre Stammzellplastizität aufrechterhalten wird.

Table of Contents

Declaration	III
Abstract	IV
Zusammenfassung	V
1 Introduction	1
1.1 How do cells specialize functionally?	1
1.2 Symmetry breaking in early mouse embryogenesis	1
1.3 Signaling control of early cell differentiation	3
1.3.1 Mouse embryonic stem cells as a model system	3
1.3.2 Transition from naïve to primed pluripotency	4
1.3.3 Epiblast and primitive endoderm differentiation	6
1.3.4 Anterior visceral endoderm differentiation	7
1.4 Signaling-induced transcriptional regulation during early differentiation	8
1.4.1 FGF signaling	8
1.4.2 Wnt signaling	10
1.4.3 mTOR signaling	11
1.5 Transcriptional regulation in mESCs	12
1.5.1 Mediator links signaling and transcriptional regulation	12
1.5.2 Mediator kinase module regulates pluripotency	15
1.5.3 Transcriptional bursting contributes to heterogeneity	16
1.6 Objectives	18
1.6.1 Quantify stochastic expression of FGF target genes	19
1.6.2 Identify molecular mediators of transcriptional changes in mESCs	19
1.6.3 Understand the role of <i>Med12</i> in transcriptional regulation during early differentiation	19
2 Materials	20
2.1 Used oligo sequences	20
2.2 Commercial kits	22
3 Methods	23
3.1 Cell culture	23
3.2 Cell lines	23
3.3 sgRNA cloning and generation of single-gene mutant mESCs	24
3.4 Genome-wide CRISPR Screen	25
3.5 Immunostaining	26
3.6 Immunoblotting	26
3.7 In situ HCR	27

3.8 Fluorescence microscopy and image analysis	27
3.8.1 Image acquisition	27
3.8.2 Image analysis	28
3.9 Flow cytometry	29
3.10 Clonogenicity Assay	29
3.11 Bulk RNA sequencing	30
3.11.1 Experimental implementation	30
3.11.2 Bulk RNA sequencing data analysis	30
3.12 Single-cell RNA sequencing experimental procedure and analysis	31
3.12.1 Droplet-based single-cell RNA sequencing and burst parameter estimation	31
3.12.2 Multiplexed single-cell RNA sequencing	32
3.13 Data and code availability	34
4 Results	35
<hr/>	
4.1 Transcriptional responses to FGF signaling in mESCs	35
4.1.1 FGF signaling induces strong transcriptional changes	35
4.1.2 Transcriptional dynamics of FGF target genes	36
4.2 Signaling and transcriptional regulators of <i>Spry4</i> expression	41
4.2.1 Establishing a <i>Spry4</i> ^{H2B-Venus} reporter line	41
4.2.2 Genome-wide CRISPR knockout screen based on <i>Spry4</i> ^{H2B-Venus} reporter activity	42
4.2.3 Positive regulators of <i>Spry4</i> expression	45
4.2.4 Negative regulators of <i>Spry4</i> expression	48
4.3 <i>Med12</i> modulates gene expression independently from signaling in pluripotency	50
4.3.1 Establishment of <i>Med12</i> mutant lines	50
4.3.2 Loss of <i>Med12</i> affects gene expression in pluripotency and differentiation	51
4.3.3 <i>Med12</i> regulates gene expression independently from pluripotency-associated signals	53
4.3.4 Limited redundancy between <i>Med12</i> and <i>Med12L</i>	57
4.4 <i>Med12</i> regulates pluripotency states and transitions	58
4.4.1 Lower clonogenicity in <i>Med12</i> mutants in the presence of FGF	58
4.4.2 <i>Med12</i> mutant cells leave naïve pluripotency delayed	59
4.5 Establishing Epi, PrE, and AVE cells in <i>Med12</i>-mutant cells	61
4.5.1 Inducible GATA6 cell line to model PrE development	61
4.5.2 Transitions to Epi and PrE identities buffered against loss of <i>Med12</i>	62
4.5.3 <i>Med12</i> mutant cells can acquire AVE fate	67
4.6 <i>Med12</i> enhances plasticity during PrE differentiation	69
4.6.1 Transcriptome-wide comparison of wild type and <i>Med12</i> mutant in PrE differentiation	69
4.6.2 <i>Med12</i> affects pluripotency homogenously	70
4.6.3 Lower mRNA counts in <i>Med12</i> -mutant cells upon differentiation	72
4.6.4 <i>Med12</i> has only minor effects on cell identities	73
4.6.5 <i>Med12</i> mutants are less plastic and have lower biological noise	76
5 Discussion	78
<hr/>	
5.1 Context of regulators identified in CRISPR screen	78
5.1.1 Connection between previously described and newly identified regulators	78
5.1.2 Limitations of the CRISPR screen approach	81

5.2 Functional characterization of <i>Med12</i>	82
5.2.1 Role of <i>Med12</i> in pluripotency	82
5.2.2 <i>Med12</i> cooperates with signaling during pluripotency transitions	83
5.2.3 mESC populations buffer PrE differentiation against loss of <i>Med12</i>	85
5.2.4 <i>Med12</i> contributes to plasticity during differentiation	86
5.3 FGF and <i>Med12</i> regulate noise during differentiation	87
5.4 Conclusions and future perspectives	88
6 References	90
<hr/>	
7 Appendix	102
7.1 List of used abbreviations	102
7.2 List of Figures	103
7.3 List of Supplementary Figures	105
7.4 List of Tables	105
7.5 List of Supplementary Tables	105
7.6 Supplementary Figures	106
7.6.1 gRNA representation over time during CRISPR knockout screen	106
7.6.2 Single-cell RNA sequencing quality control	111
7.7 Supplementary Tables	113
7.7.1 Single-cell RNA sequencing quality control	113
Acknowledgments	115

1 Introduction

1.1 How do cells specialize functionally?

Multicellular organisms consist of groups of cells that fulfill specialized functions. How do cells acquire these different identities? With the finding of mitosis, it has been proposed that all somatic cells of an organism contain the same genome. Later, nuclei from intestinal epithelium cells from adult frogs were transplanted into frog egg cells. These gave rise to a whole organism, confirming that specialized cells still contain all the genetic information required for the generation of all other cell identities (Gurdon, 1962). Instead of changing the genomic composition between different cell types, vastly different functionality is a consequence of expressing different sets of genes. Differential gene expression is a process regulated on at least four different levels: Transcriptionally, during RNA processing, translationally, and post-translationally. The establishment of new functionalities during transitions from one cell state to another, for example, during development, induces transcriptional changes as the most fundamental step of gene regulation. Therefore, the formation of new cell types can often be followed by assessing RNA production and abundance on a genome-wide scale.

An impressive example constitutes mammalian development, during which many cell types derive from a single totipotent cell type in a very fast and robust manner. Cell-cell communication within the embryo and between maternal tissues and the embryo coordinates the loss of symmetry in homogenous cell populations, referred to as a symmetry-breaking event. Reducing symmetry during development happens on multiple levels, which are inherently connected. Initially, homogenous populations of cells diverge into multiple committed cell types, constituting a symmetry-breaking event on the cellular identity level. The definition of the embryonic axes additionally requires spatial separation of these cell types. Individual cells have to maintain their capability to adapt to their changing signaling environment as long as necessary, while on the other hand mounting complex transcriptional responses. During symmetry-breaking events, cellular plasticity has to be well balanced with lineage commitment.

1.2 Symmetry breaking in early mouse embryogenesis

To further understand how cell-cell signaling leads to transitions from one cell identity to another, I focus on symmetry-breaking events very early in development. These include examples on the cellular level, and further define the main spatial axes of the embryo. Manifesting the start of exponentially increasing complexity, the first symmetry breaking-events have to be outstandingly

robust on the level of the established cell identities and their proportions, in their timing, and against environmental perturbations. To explore the mechanisms of this robustness, mouse embryos became a well-studied paradigm for early embryogenesis, due to their accessibility, limited size, and number of different cell types (Figure 1).

After initiation of development by fertilization, three rounds of cleavages give rise to a fully symmetric 8-cell state embryo. These cells, called blastomeres, are still totipotent since they can successfully contribute to all embryonic and extraembryonic tissues in chimeras (Kelly, 1977). At this point, the embryo undergoes compaction and polarization. Polarized cells divide in part asymmetrically forming a morula with inner and outer cells at day 2.75 to 3.5 after fertilization (Chazaud & Yamanaka, 2016; Saiz & Plusa, 2013). The inner cells give rise to the inner cell mass (ICM), while the outer cells form the trophectoderm (TE, Figure 1). For the first time, differentially expressed genes can be detected between these cell types: The TE marker gene *Cdx2* transcriptionally represses the expression of the ICM-specific genes *Nanog* and *Pouf5f1* (*Oct4*) in TE cells (Strumpf et al., 2005). *Cdx2* expression in the TE itself is determined by the apical localization of its mRNA in the polarized cells during the asymmetric divisions and additionally influenced by Hippo/YAP signaling (reviewed by Chazaud & Yamanaka, 2016).

Subsequently, a fluid-filled cavity is formed by TE cells secreting fluid into a forming blastocoel, positioning the ICM cells to one side of the surrounding TE cells at embryonic day 3.25 (Figure 1). This process marks a symmetry breaking event on the spatial level, predetermining the first axis of the embryo, later referenced to as the proximal-distal axis. Following this, a second bifurcation occurs within the next 12 h in the ICM, giving rise to epiblast (Epi) cells and primitive endoderm (PrE) cells (reviewed by Simon et al., 2018). These fates mark a major split of developmental trajectories since the PrE cells mostly form extraembryonic tissues contributing to the yolk sack and guiding the Epi cells' development, which form the embryo proper. Epi and PrE cell establishment is characterized by a mutually exclusive expression of PrE marking *GATA6* and the Epi marker *NANOG*. Initially, zygotic co-expression of both transcription factors starts around the 4-cell state and accumulates in all cells to the 32-cell stage. Mutual exclusive expression of the Epi and PrE cell type markers initiates in a salt and pepper-like pattern, but cell sorting and epithelialization of the PrE cells rearrange the Epi cells to lay in between the TE cells on one side and a PrE epithelium on the other side facing the blastocoel.

After sorting the PrE cells around the Epi cells, a mouse blastocyst undergoes rearrangements during uterine implantation (Figure 1). It elongates and the Epi forms a polarized epithelium, forming a second cavity. Epi cells are still surrounded by a single layer of PrE cells which progress into visceral endoderm cells. Within this cell type the anterior visceral endoderm (AVE) forms, determining the head-to-tail axis in the underlying epiblast (Figure 1). Extraembryonic tissues on

top of the cup-shaped embryo function as an important signaling center. In between the initialization of the embryonic Epi fate and the body axis determination by the AVE at embryonic day 5.5, the Epi cells undergo further pluripotency transitions. Naïve pluripotency describes an ICM-like to early Epi state, transitioning to a formative pluripotent epiblast in the cup-shaped embryo (reviewed by Smith, 2017). In the embryo, it is unclear if this more primed state is a stable cell identity or rather a transitioning phase for further differentiation (reviewed by Rathjen, 2014). Subsequently, the primed Epi cells form cells from all three germ layers in dependence on the AVE activity, marking the loss of pluripotency. Gastrulation starts, going along with major reorganizations of the embryo establishing all three germ layers ectoderm, mesoderm, and endoderm. These processes of the first 6.5 days of mouse embryonic development include important symmetry-breaking events and the definition of all major body axes resulting in a multitude of transcriptional states and cell identities.

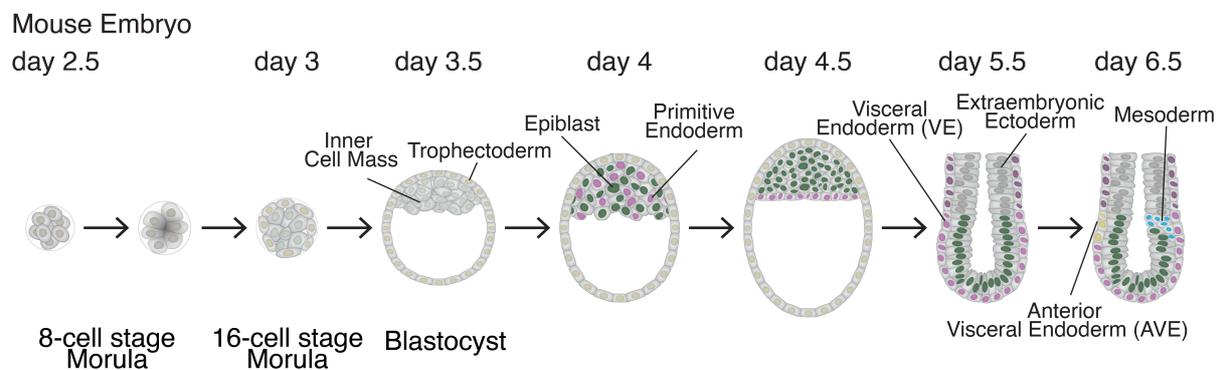


Figure 1: Early mouse embryonic development.

Schematic of cross-sections of early mouse embryonic development covering stages between 2.5 and 6.5 days after fertilization. On days 5.5 and 6.5, most extraembryonic tissues are not depicted for clarity.

1.3 Signaling control of early cell differentiation

1.3.1 Mouse embryonic stem cells as a model system

While many studies progressing our understanding of mouse embryonic development were performed *in vivo* or with embryos cultured *ex vivo*, an additive, more accessible system was established during the past four decades: mouse embryonic stem cells (mESCs; Evans & Kaufman, 1981; G. R. Martin, 1981). mESCs are derived from inner cells of peri-implantation mouse blastocysts and selected based on their capability to maintain pluripotency and proliferate indefinitely (Hooper et al., 1987). Given the correct environmental cues, mESCs have the ability to contribute to all embryonic lineages upon injection into blastocysts, thereby proving the maintenance of pluripotency during their culture. Due to their high developmental potential,

mESCs have been a widely used model system to mimic early symmetry-breaking events during early embryogenesis.

Compared to mouse embryos, mESCs are easy to maintain in large numbers, opening the possibility for large-scale genetic perturbation experiments to identify molecules regulating the cellular symmetry breaking events. Early approaches used libraries of small interfering RNAs (siRNAs) in an arrayed format to identify proteins and associated signaling systems during pluripotency and differentiation (Hu et al., 2009; S. H. Yang et al., 2012). RNA interference uses lentiviral integration of the siRNAs into the genome of the targeted cells. The expression of the siRNA mediates the cleavage of the targeted mRNAs, thereby reducing corresponding protein abundance. These perturbations were combined with phenotypic readouts, e. g. marker gene expression and cell sorting or cell survival in defined media conditions. CRISPR/Cas9 extended the possibilities for screens by the usage of guide RNA (gRNA) libraries, which together with the Cas9 activity mediate double strand breaks at targeted genes, efficiently causing gene knockouts. Together with the establishment of pooled libraries and next generation sequencing, this improved genome-wide screening, leading to the identification of essential genes for proliferation (Seruggia et al., 2019), pluripotency maintenance (M. Li et al., 2018) and differentiation (Betschinger et al., 2013; Villegas et al., 2019; B. Yang et al., 2020). Lastly, screening technologies in mESCs were extended to the activation of targeted genes with CRISPR activation (CRISPRa), identifying driver genes of differentiation (Y. Liu et al., 2018). Here, a kinase dead variant of Cas9 is targeted to genes of interest by a gRNA library together with a transcriptional activator, causing the expression of the targeted gene.

Utilizing the signaling molecules identified for example in genetic screens, mESCs allow precise control changes in the signaling environment, and symmetry-breaking events can be mimicked in various 2D and 3D systems (Baillie-Benson et al., 2020). Recently, mESC descendant Epi and PrE cells have been included in 3D models progressing to stages corresponding to embryonic day 8.5 (Amadei et al., 2022; Tarazi et al., 2022), emphasizing the self-organizing potential during embryonic development.

1.3.2 Transition from naïve to primed pluripotency

The transition from naïve pluripotency to a more primed state and finally lineage commitment in the embryonic epiblast between embryonic day 3.5 and 6.5 is difficult to study *in vivo* given its position in the center of the embryo. mESCs helped to understand this important step preparing the formation of the three germ layers. mESCs can proliferate in a naïve pluripotent state in a culture in two major media conditions: Based on Glasgow Minimum Essential Medium (GMEM)

in presence of fetal bovine serum (ES + LIF) or in chemically defined N2B27 supplemented with PD0325901 (PD03), CHIR99201 (Chiron, Ying et al., 2008). The small molecule inhibitors PD03 and Chiron reflect the importance of Fibroblast growth factor (FGF) and *Wingless iNTEgrated* (Wnt) signaling during pluripotency. PD03 inhibits FGF signaling and ERK activity in mESCs (Bain et al., 2007; Barrett et al., 2008). Chiron acts as a Wnt signaling agonist, leading to global Wnt activity (Cline et al., 2002; Ring et al., 2003). Additionally, LIF is used in the presence of PD03 and Chiron to stabilize pluripotency further (Niwa et al., 2009). The signaling environment in ES + LIF and N2B27 + PD03 + Chiron + LIF (2i + LIF) are contradicting since paracrine FGF signaling establishes high ERK levels in ES + LIF medium and Wnt signaling levels are low due to the absence of an activator. However, the presence of LIF and other factors from the serum keep mESCs pluripotent, although with a higher heterogeneity in the expression of naïve marker genes (Morgani et al., 2017). In both cases the naïve state pluripotency is characterized by the unique expression of a set of transcription factors, centered around OCT4, NANOG, and SOX2 and by a hypomethylated genome. Moreover, mESCs cultured in either media condition can contribute to all embryonic lineages of an embryo, while culturing in 2i + LIF additionally preserves low-level ability to contribute to extraembryonic lineages (Morgani et al., 2013; Nichols et al., 1990).

In vivo, naïve pluripotency corresponds to an early Epi state, which transits into a primed pluripotent state in later Epi cells. With the help of mESCs this transition was analyzed with greater details and time resolution (Figure 2). Naïve pluripotent cells first transit into a rosette-stage pluripotency induced by the downregulation of Wnt signaling (Neagu et al., 2020). This rosette stage corresponds to a transient state in vivo just when the epiblast starts to form a lumen. The naïve pluripotency marker KLF4 is downregulated, while OTX2 starts to be expressed. In mESCs, this transition is reversible by activating Wnt signaling. Further progression and commitment on the route to a primed state passing a formative state is mediated via activation of FGF/ERK signaling (Kalkan et al., 2019).

In addition to FGF and Wnt signaling, multiple other signaling systems have been proposed to influence the naïve to primed transition in mESCs. Mutation of *Rbpj*, a Notch signaling effector, efficiently prevented exit from pluripotency in combination with mutations of *Tcf3* and *Etv5* (Kalkan et al., 2019). Utilizing genomic screens, pluripotency was described to be stabilized upon perturbation of mammalian target of rapamycin (mTOR) signaling-related genes *Flcn* and *Tsc2* (Betschinger et al., 2013). In summary, knowledge about signaling systems involved in maintaining and priming of pluripotency was confirmed and extended using mESCs. However, it remains speculative how this multitude of signals cooperates with the general transcriptional machinery at pluripotency and developmental genes.

1.3.3 Epiblast and primitive endoderm differentiation

The inner cell mass divides into Epi and PrE cells at embryonic day 4, marking a predominantly signaling-induced symmetry-breaking event in the embryo. Mutual exclusion in this case goes along with mutual inhibition between GATA6 and NANOG. *Gata6*-mutant embryos show NANOG expression in all ICM cells (Bessonnard et al., 2014; Schrode et al., 2014), while *Nanog*-mutant embryos maintain a GATA6 positive ICM (Frankenberg et al., 2011). Mutual expression of the transcription factors NANOG and GATA is crucial for establishing the two cell types. Although the loss of one of the markers GATA6 or NANOG predicts differentiation outcome early, cells maintain their plasticity until embryonic day 4.0, after mutually exclusive expression is established (Yamanaka et al., 2010). Especially for the PrE cells, this plasticity to change fates can extend until embryonic day 4.5 (Grabarek et al., 2012). This plastic behavior was confirmed in vitro. PrE-cells, sorted after 20 h of differentiation for the expression of a GATA6 reporter, were able to reestablish both Epi and PrE cell types in N2B27 (Raina et al., 2021).

The ability of cells to change fate late during their trajectory to acquire another cell identity indicates coordination via cell-cell signaling. Indeed, *Fgf4* is crucial for differentiation progression from the ICM. *Nanog* mutant embryos lack expression of *Fgf4* (Frankenberg et al., 2011), usually marking cells differentiating into Epi cells (Frankenberg et al., 2011; Guo et al., 2010; Ohnishi et al., 2013). Vice versa, altering *Fgf4* signaling levels affect PrE and Epi proportions: High levels of FGF4 promote PrE differentiation, while downstream inhibition of FGF-signaling impairs PrE formation (Yamanaka et al., 2010). *Fgf4*-mutant embryos expressed GATA6, indicating that FGF4 is not required for the initial upregulation of GATA6 in the embryo, but did not form mutually exclusive GATA6 and NANOG expression patterns in the ICM (Kang et al., 2013; Krawchuk et al., 2013). In summary, establishing Epi and PrE fates from ICM cells marks a well-studied cell-cell communication-based bifurcation during early development, one of the first in which different gene expression programs are activated, centered around the mutual inhibition of the transcription factors GATA6 and NANOG.

The knowledge of the dependency on first GATA and NANOG co-expression and then on the presence of FGF signaling helped to mimic this lineage bifurcation in vitro with mESCs (Figure 2). In ground state pluripotency, mESCs are distinct from the in vivo ICM-state by the lack of expression of GATA factors. Inducing a transgene coding for this transcription factor and subsequent differentiation lead to primitive endoderm differentiation (Fujikura et al., 2002; Schröter et al., 2015; Wamaitha et al., 2015). The established proportions of both Epi and PrE cells are similar to the embryo (Plusa et al., 2008; Raina et al., 2021). Confirming embryonic studies, a crucial role for paracrine FGF signaling was described in mESCs: The level of FGF/ERK signal sets the threshold of GATA levels needed for PrE differentiation. On a population level, cells control

their cell type ratio by adjusting local FGF signaling. This ensures the robustness of cell type proportioning to diverse starting conditions with respect to GATA induction levels (Raina et al., 2021; Schröter et al., 2015). In contrast, titration of FGF4 in *Fgf4*-mutant cells during differentiation increases the ratio of PrE cells (Raina et al., 2021), recapitulating the phenotypes *in vivo*.

1.3.4 Anterior visceral endoderm differentiation

Both Epi and PrE are not only marked by specific set of marker genes, respectively, but can also develop or contribute to self-organized 3D structures in multiple embryo-like aggregate systems (Amadei et al., 2022; Schumacher et al., 2023). Following the establishment of PrE and Epi identities and major rearrangements, the anterior-posterior axis as the next major body axis of the embryo is laid down by a symmetry breaking event. First, cellular symmetry is broken by the formation of a specialized cell population at the distal tip of the embryo within the VE layer, characterized by the expression of the transcription factors *Otx2* and *Eomes* (Hoshino et al., 2015; Nowotschin et al., 2013). The anterior-posterior axis is finally determined by the position this AVE population migrates towards. Although this cell population is thought to be locally restricted by an external BMP gradient from extraembryonic tissues, bilayered embryonic aggregates lacking these extraembryonic tissues show the potential to form and locally restrict an AVE-population, derived from PrE cells. Furthermore, a 2D layer of PrE cells forms nests of AVE cells within a signaling environment mimicking the contact with Epi cells (Figure 2; Schumacher et al., 2023). This happens in dependence of β -Catenin, as tuning β -Catenin activity manipulates the proportion of cells expressing AVE marker genes (Schumacher et al., 2023). This further proves the functionality of *in vitro* differentiated PrE cells. In summary, the exit of pluripotency leading to Epi-like cells, as well as both the bifurcation into Epi and PrE cells from induced mESCs and further differentiation into an AVE population, recapitulates the corresponding *in vivo* processes and the involved signaling systems.

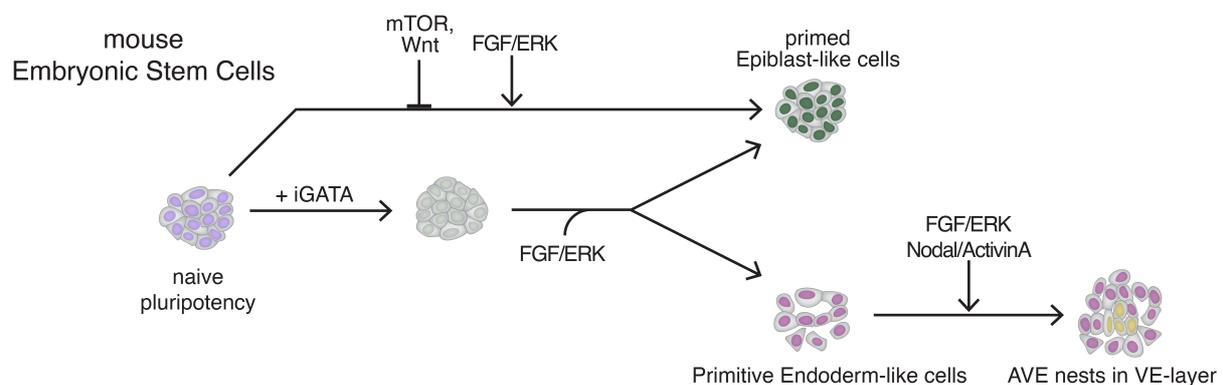


Figure 2: mESCs recapitulate symmetry breaking in early development.

(Continued on next page)

Schematic of mESC differentiation from naïve pluripotency to primed epiblast-like cells or via GATA induction and paracrine FGF/ERK signaling into mixtures of epiblast-like and primitive endoderm-like cells.

1.4 Signaling-induced transcriptional regulation during early differentiation

To coordinate the symmetry-breaking events on the cell population level, including formation of mixtures of PrE and Epi cells, and on the spatial level, for example the formation of the locally restricted AVE population, cells communicate about their fate and relative position in the embryo to their surroundings. Cell communication is a recursive process during which cells react to their surrounding and this reaction itself again modifies the signaling environment of their own and of neighboring cells. In order to react to their environments, are sensing the biophysical properties of their surroundings. Additionally, the reception of biochemical signals potentiates the possibilities for communication by the plethora of different signaling molecules. Together with the signaling dynamics, the signaling molecules define, in which range and quantity a signal is received by cells. By limiting this work to the described differentiation events during early embryonic development the diversity of biochemical signals involved is more defined. For example, the establishment of Epi and PrE cells from the ICM was found to be dependent on FGF signaling. As described above, Wnt- and mTOR-dependent signaling states are crucial for the exit of a naïvely pluripotent state of the Epi cells. The following paragraphs will first focus on one of these signaling systems at a time, simplifying how cells context-dependent integrate various signals. Many of the information about signaling in early embryogenesis introduced below were either first obtained *in vivo* and afterward tested in more detail in mESCs or vice versa. To understand how the extracellular signals lead to changes in of the cell's state, the description of the triggered intracellular events marks a major step. Importantly, this includes how the different signals lead to transcriptional regulation and how the respective target genes are defined.

1.4.1 FGF signaling

FGF-dependent cell-cell communication during embryonic development is based on the secretion of extracellular proteins from the FGF family with 22 members. Most subgroups of these signaling molecules function as autocrine or paracrine. Heparin sulfate proteoglycan molecules bind to FGFs thereby limiting their diffusion through the extracellular matrix and promoting binding to FGF receptors (FGFRs), which induces their dimerization (Figure 3; Matsuo & Kimura-Yoshida, 2013; Yayon et al., 1991). The intracellular domain of the FGFRs trans-autophosphorylate several tyrosine residues in the cytoplasmic domain of the receptor itself, promoting the binding of further effector proteins. Binding of FGF to one of four FGFRs can cause the activation of multiple

intracellular signaling cascades, including JAK/STAT (Janus kinases (JAK) - signal transducer and activator of transcription proteins (STAT)) signaling, Phosphoinositide phospholipase C (PLC) signaling, protein kinase B (AKT) signaling and Ras GTPase- (from **Rat sarcoma virus**) mediated extracellular signal-regulated kinase (ERK) activation. In preimplantation development, the effects of FGF signaling have been predominantly linked to the RAS-dependent activation of ERK (Nichols et al., 2009; Stavridis et al., 2007). This cascade is triggered by the binding of FGFR substrate 2 (FRS2) to the activated intracellular juxtamembrane region of FGFRs by the interaction with its phosphotyrosine-binding domains. Growth Factor receptor Bound 2 (GRB2) and Son Of Sevenless (SOS) are recruited, exchanging the nucleotide bound to Ras from GDP to GTP. GTP-bound Ras activates the Raf kinase (from **Rapidly Accelerated Fibrosarcoma**) by phosphorylation which itself activates the mitogen-activated protein kinase kinase (MEK), propagating and amplifying the phosphorylation status to ERK. PD03, which as described above contributes to stabilizing pluripotency in mESCs, specifically inhibits MEK and prevents downstream ERK activity (Bain et al., 2007; Barrett et al., 2008). Within the FGF-signaling cascade, feedback mechanisms control the dynamics of ERK signaling on different time scales (Lake et al., 2016). Phospho-ERK activates a plethora of intracellular components and processes (Ünal et al., 2017).

FGF/ERK signaling is central to coordinate differentiation of multiple cell types in early mammalian development. Specifically, pharmacological inhibition, mutagenesis of both FGFRs as well as mutagenesis of only FGF4 prevents PrE differentiation in the preimplantation stage (Kang et al., 2013, 2017; Molotkov et al., 2017; Nichols et al., 2009), emphasizing that only FGF4 is essential for FGF-signaling for ICM bifurcation. Additionally, the epiblast cells also depend on active FGF4 signaling to mature from the ICM (Ohnishi et al., 2013).

For these FGF-dependent lineage transitions, transcriptional regulation of FGF-specific target genes is one of the ERK-regulated processes. Known FGF target genes in early mouse development include *Spry2* and *4* as well as *Dusp4* and *6* (Morgani et al., 2018). A transcriptional reporter for *Spry4* was used to monitor FGF/ERK signaling during developmental progression in the mouse embryo and mESCs (Morgani et al., 2018). How FGF/ERK signaling influences gene expression in these differentiation contexts is elusive. In general, ETS domain-containing transcription factors were identified to act downstream of FGF signaling (Ornitz & Itoh, 2015; Sharrocks, 2001). Potential sequence-specific transcriptional factors activated by ERK in mESCs include ETV4 and ETV5. Mutation of both genes however only partly phenocopied the inhibition or mutation of FGF signaling (Akagi et al., 2015; Kalkan et al., 2019). ERK signaling furthermore controls the activity of the chromatin modifier polycomb repressive complex 2 (PRC2) (Tee et al., 2014) and potentially translocates to the nucleus itself directly interacting with RNA polymerase II (Göke et

al., 2013; Tee et al., 2014). The specific effect of such general regulation implies a specific composition of the transcriptional complex at developmentally important genes, offering mechanisms of regulation beyond the specificity of transcription factors.

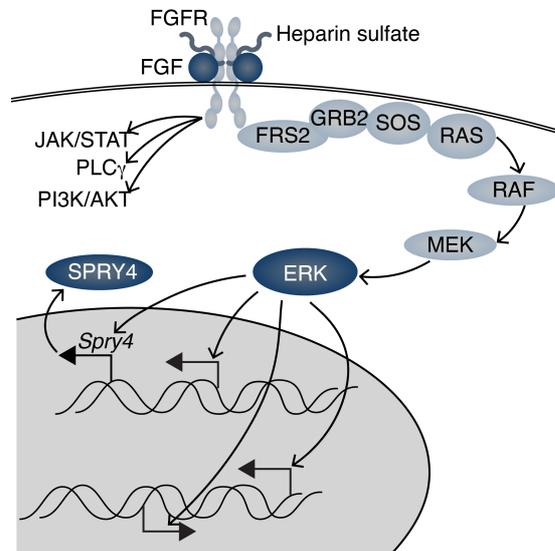


Figure 3: Schematic of FGF-signaling cascade in mouse embryonic stem cells.

FGF binding triggers an intracellular cascade predominantly leading to ERK activation. During differentiation, ppERK changes the transcription of so far only partly described FGF-specific target genes, including *Spry4*.

1.4.2 Wnt signaling

Wnt signaling is crucial for multiple steps in early embryo development. The Wnt family comprises 11 conserved members among vertebrates, with most mammalian genomes entailing 19 WNT genes in 12 conserved subfamilies (reviewed by Clevers & Nusse, 2012). WNT proteins, approximately 40 kDa in size, undergo modification in the endoplasmic reticulum through lipid addition, crucial for their secretion and effective signaling. Canonical WNT signaling involves the interaction between WNT proteins and a receptor complex comprising Frizzled and LRP5/6, inhibiting part of the β -catenin degradation complex. This inhibition allows β -catenin to translocate into the nucleus. Reaching the nucleus, β -catenin binds and activates the TCF/LEF transcription factors, leading to transcription of Wnt-responsive genes. In the absence of WNT, β -catenin degradation occurs via a destruction complex involving DVL, APC, AXIN, and GSK3. The small molecule inhibitor Chiron functions as a GSK3 inhibitor, which ligand independently activates β -Catenin activity (Cline et al., 2002; Ring et al., 2003). Non-canonical WNT signaling pathways influence cytoskeletal rearrangements, calcium release, and cellular responses mediated by different receptors and intracellular proteins (Nusse, 2012). Specifically, in the planar cell polarity (PCP) pathway, certain Wnt ligands bind a receptor consisting of Frizzled and

Ror instead of LRP5/6 and activate Disheveled in a way allowing it to interact with RhoGTPases. These remodel the cytoskeleton and thereby influence cell shape and movement.

During early embryonic development, Wnt ligands have been detected to be expressed distinctly regionalized in different tissues, implying specific roles for multiple differentiation directions (reviewed by Tepekoy et al., 2015). In the blastocyst, higher Wnt activity is detected in the epiblast, compared to PrE cells. This was linked to the activity of the transcriptional repressor TCFL1 acting on pluripotency genes (Athanasouli et al., 2023). Additionally, Wnt signaling represses the transcription factor TCF3 activity, so that naïve pluripotency transcription factor genes, for example, *Esrrb* and *Nanog*, are no longer repressed (Martello et al., 2012; Pereira et al., 2006). Mutation of the destruction complex subunit APC in the mouse embryo leads to activation of β -catenin and its transfer to the nucleus preventing the induction of the AVE (Chazaud & Rossant, 2006). In summary, Wnt signaling is a signaling system based on a multitude of ligands emphasizing its context-dependent and dynamic functions implemented by gene specific activities of TCF/LEF transcription factors. Some of the transcription factors and Wnt signaling molecules are essential in early embryonic stages.

1.4.3 mTOR signaling

Another central complex regulated by extracellular signals is mTOR. In mammals, two mTOR complexes (mTORC1 and mTORC2) can be distinguished by their composition. Together, they fulfill diverse functions in many cell types regulating cellular and physiological homeostasis. Specifically, mTORC2 for example regulates apoptosis, whereas mTORC1 is associated with the regulation of autophagy, lipid synthesis, mitochondrial metabolism, and mitochondria biogenesis (reviewed by Liu & Sabatini, 2020). mTORC1 positively regulates protein biosynthesis by increasing both transcriptional and translational activity. Ribosome biogenesis is enhanced by inducing the transcription of ribosomal RNAs (Mayer et al., 2004). These diverse effects of mTOR signaling are coordinated by a plethora of integrated signals. Major signals include growth factors in the surroundings, the energy status of a cell, oxygen levels, and amino acid abundance. Many of these signals influence the activity of the tuberous sclerosis complex (TSC), a heterodimeric complex formed from TSC1 and TSC2. This complex negatively regulates mTOR activity by functioning as a GTPase activating protein for the GTPase RHEB, which only stimulates mTOR in its active GTP-bound form (Long et al., 2005). Growth factors act on the activity on TSC in opposing ways: While growth factor-induced active RAS signaling results in ERK activity, which phosphorylates TSC leading to inactivation (Ma et al., 2005), Wnt signaling via GSK3 phosphorylating TSC leads to activation of the complex (Inoki et al., 2006).

While these early studies were not performed in the context of early embryonic development, the intertwined nature with multiple signaling systems, cell growth regulation, and metabolism suggest a crucial role during development. Indeed, homozygous mutation of mTOR components leads to early growth arrests in mouse embryos, indicating the impairment of mTOR-deficient Epi cells to exit pluripotency (Gangloff et al., 2004). Additionally, mouse embryos can reversibly enter a diapause state, in which transcription is globally suppressed by a reduced mTOR activity within the pluripotent cells. This state is triggered by unfavorable conditions *in vivo* and can be pharmacologically induced by partially inhibiting mTOR (Bulut-Karslioglu et al., 2016). Genomic screens revealed that transcriptional regulation induced by mTOR is linked to *Tfe3* in the epiblast. mTOR activity regularizes the subcellular localization of Tfe3 and thereby connects differentiation with catabolic processes in mESCs (Betschinger et al., 2013; Villegas et al., 2019), in line with regulating diapause in dependence on the nutrients provisioning.

1.5 Transcriptional regulation in mESCs

The introduced signaling systems in mESCs discussed above lead to transcriptional changes in order to establish new functionalities in cells. The signal has to be linked to a change of RNA polymerase II (PolII) activity at specific target genes. This can be mediated by a sequence-specific transcription factor, which is activated by the intracellular signaling cascade and binds target genes based on DNA motifs in target promoter and enhancer regions. In mESCs, these sequence specific transcription factors include TCF/LEFs for Wnt signaling and TFE3 downstream of mTOR. However, the specific transcription factors usually do not influence PolII directly. Co-activators bind to the sequence-specific transcription factors and implement changes of transcriptional activity at target promoters. Moreover, signaling cascades can also directly modify these co-activators independent of DNA motif specificity. Regulating the core transcription machinery potentially influences transcription globally. How these effects can lead to signaling specific transcriptional responses and are coordinated with the sequence-specific activity of transcription factors so far remained an open question.

1.5.1 Mediator links signaling and transcriptional regulation

Specific DNA binding motifs for transcription factors are often organized in enhancer regions, which can be located megabases away from the transcriptional start site of the respective gene. An important co-activator for the transfer of information about transcription factor binding activity to the core promoter and directly to the protein-coding gene transcribing RNA polymerase II (PolII) is the Mediator complex. Since sequence-specific transcription factors do

not recruit the polymerase themselves and Mediator physically bridges enhancers and core promoter regions, this complex potentially serves as an important integrator of signaling and mediates transcription factor regulation to transcriptional responses. In general, Mediator is evolutionarily conserved, comprising 25 subunits in yeast and up to 30 subunits in higher mammals (Soutourina, 2018). These subunits are organized into modules: the head, middle, and tail modules constitute the core mediator, while the mediator kinase module transiently interacts with it. The presence of Mediator is required for transcription in yeast and mammals, however individual subunits are only essential in specific contexts and cell types. The complex influences transcription in multiple ways, with each module fulfilling individual functions (Figure 4). Key to initiate transcriptional regulation by Mediator is the recruitment of subunits of the tail module, together with other transcriptional co-activators, by specific transcription factors at enhancer regions (Figure 4A,B; Grünberg et al., 2016). Next, the whole mediator complex including the Mediator kinase module interacts with cohesion to stabilize the formation of a genomic loop (Kagey et al., 2010), bringing the enhancer in proximity to the core promoter and the transcriptional start site (Figure 4B). After dissociation of the kinase module, especially the middle and head modules of Mediator are required to recruit the preinitiation complex (PIC) consisting additionally to Mediator out of transcription factor IIA (TFIIA), TFIIB, TFIID, TFII E, TFII F, TFII H, and PolII (Figure 4C). Mediator directly interacts with the C-terminal domain (CTD) of the Polymerase (Näär et al., 2002). This CTD contains repeats that get phosphorylated by TFII H. Phosphorylation of enough repeats causes release of PolII from Mediator and the start of transcription (Corden, 2013). Even after transcription initiation Mediator takes part in regulating the release from promoter-proximal pausing and therefore efficient transcriptional elongation. During promoter-proximal pausing PolII accumulates downstream of the transcriptional start site, which might serve as a mechanism of priming genes for immediate activation contributing to the reactivity and plasticity on the level of single promoters. Pause release is positively influenced via binding of the Mediator kinase module (Figure 4D; Donner et al., 2010; Galbraith et al., 2013). Moreover, Mediator potentially evicts nucleosomes from regions just around the promoter to allow the transcriptional complex to form (Kremer et al., 2012).

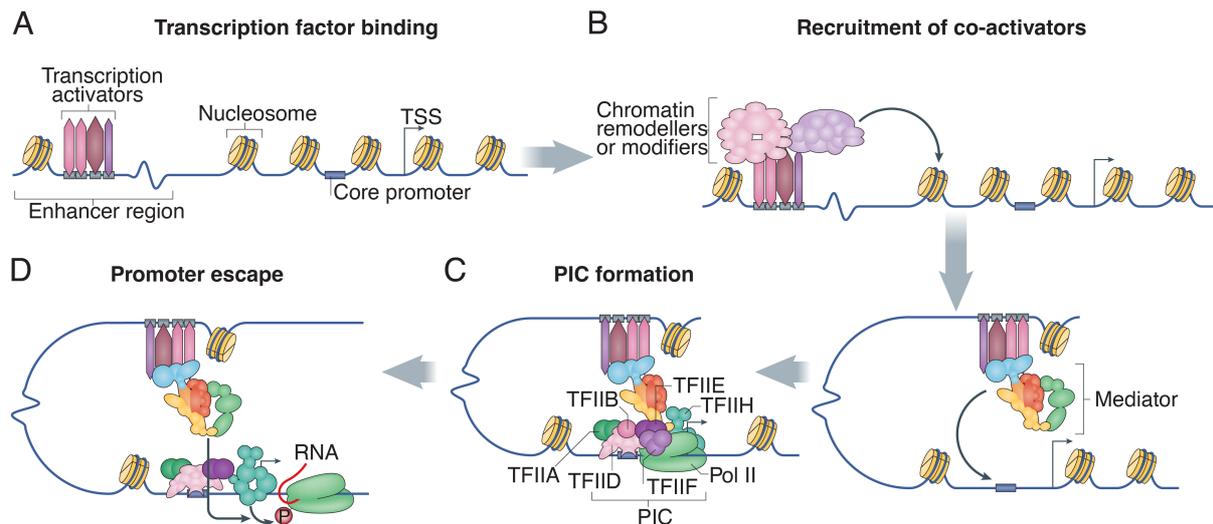


Figure 4: Mediator links enhancers and promoters to regulate transcription.

Reproduced from Soutourina, 2018 with permission from Springer Nature.

A Activation of transcription starts with the binding of transcriptional activators to enhancer regions.

B Co-Activators are recruited and chromatin loop is established.

C Formation of the Preinitiation complex (PIC), requiring the dissociation of the mediator kinase module (green).

D Start of transcription and elongation, regulated by the Mediator dependent phosphorylation via TFIIH.

Within a defined cell state, this multitude of mechanisms of how Mediator regulates transcription are partly specific for characteristics of the individual promoters. For example, the effects of Mediator depend on the chromatin surrounding of the transcribed gene: Promoters containing TATA-boxes have higher nucleosome dynamics, correlating with the higher dependency on Mediator to establish the PIC. This correlation might be underlying the higher plasticity and gene expression flexibility of TATA-box-containing promoters (Eychenne et al., 2016). Generally, cell identity-defining genes are often characterized by unusually strong corresponding enhancer elements, referred to as super-enhancers, which are characterized by dense occupation with Mediator complexes. In pluripotent mESCs, these genes include the key pluripotency transcription factors OCT4, SOX2, and NANOG (Whyte et al., 2013).

Furthermore, the gene specificity of Mediator is not only dependent on fixed promoter characteristics, but also actively regulated during cell type transitions. The interaction of cohesion with Mediator is limited to individual sets of genes in mESCs and mouse embryonic fibroblasts, respectively (Kagey et al., 2010). Cell-type specificities like this were previously linked to the cooperation of individual Mediator subunits with specific signaling-regulated proteins. For example, in HaCaT (spontaneously transformed aneuploid immortal keratinocyte) cells TGF β signaling affecting SMAD2/3 activity is impaired in *Med15* knockdowns, due to a missing direct interaction between SMAD2/3 and MED15. In contrast, BMP signaling affecting SMAD1/5 is not dependent on MED15, corresponding to a lack of the interaction site in SMAD1/5 (Zhao et al.,

2013). Another transcription factor interacting with a specific Mediator subunit is the ETS domain-containing transcription factor ELK1. The transcriptional activation domain of ELK1 is phosphorylated by ERK, which controls its kinetics of the interaction with MED23 and thereby influences transcriptional output quantitatively (Mylona et al., 2016). While the deletion of *Med23* (formerly called *Sur2*) caused disruption of FGF/ERK signaling in mESCs (Stevens et al., 2002), it had only a mild effect in mouse embryonic fibroblasts, likely due to the upregulation of ELK3, which promotes transcription independent of MED23 (Balamotis et al., 2009). *Med23* and *Med15* thereby mark two examples of Mediator subunits of the tail module linked to signaling systems via their interaction with specific transcription factors and therefore context-dependent functions.

1.5.2 Mediator kinase module regulates pluripotency

The Mediator kinase module transiently binds the core mediator during recruitment of the PIC, has to dissociate during transcriptional initiation, and potentially interacts with paused transcriptional complex. Many of these processes are actively regulated, making the kinase module a node of transcriptional regulation (reviewed by Luyties & Taatjes, 2022). The kinase module consists out of four subunits: the CDK8 kinase, CCNC, MED12, and MED13. CDK8 and CCNC form a dimer executing the kinase activity. The interaction with the core Mediator complex requires MED13 (Knuesel et al., 2009; Tsai et al., 2013), while MED12 connects the CDK8-CCNC dimer with MED13 (Y. C. Li et al., 2021). In mammals CDK19, MED12L, MED13L are mutual exclusive paralogs to CDK8, MED12 and MED13, however, their contribution to the Mediator kinase module function remains elusive (Daniels et al., 2013).

One of the first signals identified acting via the kinase module is the serum response network. Specifically, CDK8-depletion leads to slower transcriptional elongation at serum-responsive genes (Donner et al., 2010). During mouse preimplantation development and in mESCs, inhibition of the CDK8/19 activity with small molecules has been described to maintain naïve pluripotency and mimic the effects of 2i treatment. Interestingly, this effect requires the physical presence of the kinase module, since the double-knockout of CDK8/19 did not maintain naïve pluripotency features of mESCs in the absence of 2i. Mechanistically, CDK8/19 inhibition was linked to the loss of a repression by Mediator of pluripotency-associated genes and enhancers (Lynch et al., 2020). The Mediator kinase module is further involved in embryonic development by MED12, which potentially acts downstream of Wnt signaling. Consistently with Wnt signaling interventions, mouse embryos do not form an AVE and did not complete gastrulation upon loss of *Med12* (Rocha et al., 2010). In addition to being essential for the structural formation of the kinase module,

Med12 might tether the kinase module to enhancer regions by its interaction with small non-coding RNAs produced in proximity to enhancers (Lai et al., 2013). This possibly keeps the kinase module at close distance to the transcriptional complex for downstream phosphorylation events during pause release and elongation. A role of *Med12* in pluripotency regulation has been suggested by work finding MED12 to interact with NANOG, which together regulate the expression of *Nanog* and *Nanog* target genes (Apostolou et al., 2013; Tutter et al., 2009). For *Med12*, a further function has been described in TGF β signaling, blocking signaling directly at the cytoplasmic side of the TGF β receptor (Huang et al., 2012). Together these studies offer multiple possibilities for how signaling and signaling-mediated transcription factor activity can lead to mediator-dependent transcriptional regulation. However, not all characteristics of Mediator and Mediator kinase-dependent transcriptional regulation are based on cell-cell signaling, but can also depend on gene characteristics like the chromatin structure. It has not been studied, how the diverse mechanism of Mediator and especially its kinase module contribute to cellular plasticity during cell type transitions of mESCs.

1.5.3 Transcriptional bursting contributes to heterogeneity

The signaling environment determines the transcriptional identity of a cell to establish multiple specialized features during differentiation. Cells within the same environment, receiving the same signals have to break symmetry to progress development. What creates the first molecular difference determining the bifurcation? Even within homogenous signaling environments, cell-to-cell variability exists, commonly referred to as noise. Noise levels can be traced back to the stochasticity of biochemical processes (intrinsic noise) and extrinsic fluctuations of cellular components (Elowitz et al., 2002). Stochasticity of events, especially in a population with a small number of cells could induce symmetry breaking. Due to the limited number of DNA repeats within each cell, stochasticity majorly influences mRNA production. Comparing the mRNA count numbers of single genes in mammalian cells revealed a high variation within cell populations (Raj et al., 2006), leading to a telegraph model (also referred to as Two-state model) of transcription. First proposed for gene product synthesis in general (Peccoud & Ycard, 1995), this model assumes a promoter existing in either an active or an inactive state (Figure 5A). Switching between the states occurs with k_{on} and k_{off} rates, while only in the active state mRNAs are produced with a gene-specific synthesis rate (k_{syn}). Relative short on- compared to off-states cause transcriptional bursts, which together with mRNA degradation (k_{deg}) lead to variability of mRNA number over time in each cell (Figure 5B, left). Due to the stochastic nature of these bursts, this explains higher intrinsic noise in mRNA count distributions within cell populations, than would be expected for a single constant synthesis rate. Mechanistically, the switch from an off state to an actively

transcribed gene includes the involvement of a plethora of protein subunits and interactions to form the transcriptional complex, including the interaction of enhancer regions with the PIC, as described above. Once all required transcription factors are in proximity, often multiple polymerases start transcription, leading to a strong burst in short time scales.

In a telegraph model of transcription, two major parameters can describe the transcription of a gene: The frequency of the occurrence of transcriptional bursts, and how many mRNAs are produced with each burst, referred to as burst size. Both parameters are potentially regulated by cell signaling. Nevertheless, they are determined not by a single master regulator but controlled by multiple factors (Ochiai et al., 2020), reflecting the multitude of involved interactions and proteins. However, burst size is described to be mainly controlled by the promoter region close to the transcriptional start site (Larsson et al., 2019), while distal enhancer regions and their interaction with the transcriptional complex regulate burst frequency (Bartman et al., 2016; Fukaya et al., 2016; C. Li et al., 2018). Different kinds of stimulation of the same gene may result in similar average expression levels, but vastly different intrinsic noise depending on if the stimulation affects burst frequency or size (Figure 5B, right). Transcriptional states of individual genes can be followed by live imaging of transcriptional activity (Bertrand et al., 1998; Hoppe et al., 2020; Ochiai et al., 2014). An experimentally more accessible approach to study transcriptional dynamics uses individual distributions of mRNA counts per cell. From these distributions, transcriptional burst parameters can be estimated in dependency of the mRNA degradation rate (Abranches et al., 2014; Raj et al., 2006, 2008). Extending the potential of this approach, the capture of mRNA counts globally with single-cell RNA sequencing and efficient estimation algorithms were established in the past decade (J. K. Kim & Marioni, 2013; Larsson et al., 2019; Ochiai et al., 2020; Vu et al., 2016).

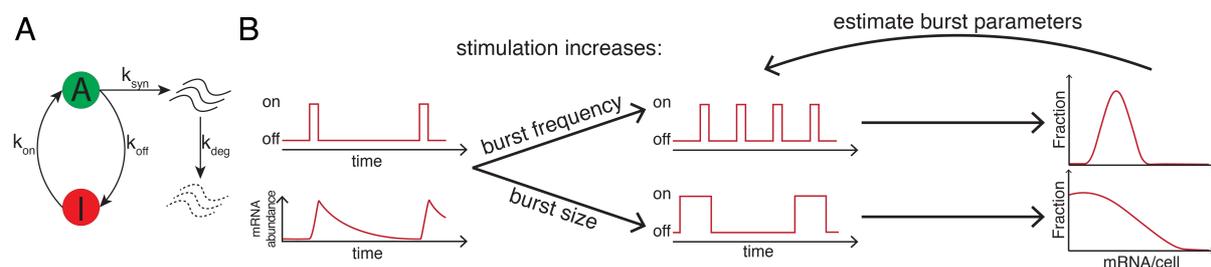


Figure 5: Telegraph model of transcriptional bursting leading to differently shaped mRNA count distributions.

A Schematic of the Telegraph model of transcription. Promoters can switch between an active (A) and inactive (I) state at the rates of k_{on} and k_{off} , respectively. Only in the active state, transcription produces mRNAs at the synthesis rate k_{syn} . mRNAs are degraded at a degradation rate of k_{deg} .

B Schematic of stimuli leading to increased burst frequency (top) or burst size (bottom) leading to similar average expression levels but differences in heterogeneity, represented by the mRNA count distributions (right).

How variability in mRNA abundance over time or within cell populations affects protein abundance and thereby functional activity, is dependent on the relation of protein turnover rates and the burst frequency. Since protein half-lives are generally longer, transcriptional bursting is at least partly buffered (Raj et al., 2006). Nevertheless, stochastic gene expression can have a functional impact on symmetry-breaking events. In the fly embryo, for example, mosaic expression of the photoreceptors for color vision is defined by a stochastic burst of *spineless* expression (Wernet et al., 2006). In mESCs, stochastic expression of the pluripotency gene translates into heterogeneity also on the protein level, shown by a high correlation between *Nanog* mRNA abundance and NANOG protein levels on single cell level (Ochiai et al., 2014). This heterogeneity allows the cells to explore pluripotency states and low NANOG expression primes cells for differentiation (Abranches et al., 2014). Additionally, in vivo, heterogenous expression of NANOG in the ICM during establishment of the Epi and PrE fate marks differentiation outcome early (Xenopoulos et al., 2015), suggesting stochastic gene expression as a driver of symmetry breaking on a cellular population level. Global analysis of bursting regulation revealed that in mESCs transcriptional bursting is regulated by the PRC2 complex as well as transcriptional elongation factors. This is potentially reflected in the observation of MED12 being enriched at promoters regulated in their burst size (Ochiai et al., 2020). Moreover, the AKT and ERK signaling state is linked to the burst kinetics by the regulation of the elongation efficiency. ERK inhibition contributes to a decreased intrinsic noise with globally lower burst sizes (Ochiai et al., 2020). The regulation of the stochastic nature of transcription appears to be important for differentiation and symmetry breaking since transcriptional bursting contributes to biological noise. How signaling systems influence the bursting and noise levels during early mammalian development and mESCs is not yet completely understood. Noisy gene expression could be an attractive possibility for cells to explore the transcriptional space on a differentiation trajectory, thereby contributing to the initiation of cell type transitions and bifurcations.

1.6 Objectives

For cell state transitions and cellular symmetry breaking, signaling has to cooperate with transcriptional regulators. Focusing on mESCs and their potential to mimic early developmental symmetry breaking events on a cell population and even spatial scale, I aimed to investigate this coordination particularly with regard to the regulation of cellular noise and plasticity. Therefore, I addressed the following aspects of transcriptional regulation in early mammalian cell differentiation.

1.6.1 Quantify stochastic expression of FGF target genes

First, my goal was to pinpoint the specific genes directly affected by FGF signaling in pluripotent mESCs and determine how sensitive each gene responds to FGF stimulation. I planned to analyze the transcription of these genes by inferring their dynamic transcriptional parameters. Therefore, I attempted to measure mRNA abundance globally with single-cell RNA sequencing. This should further the understanding of how graded transcriptional responses are established and how cell-to-cell variability in gene expression potentially influences differentiation initiation and progression.

1.6.2 Identify molecular mediators of transcriptional changes in mESCs

My next goal was to explain how the identified FGF target genes are regulated and to investigate transcriptional changes triggered by other signaling pathways relevant to development. To this end, I conducted a comprehensive CRISPR knockout screen across the genome in pluripotent mouse embryonic stem cells (mESCs). By employing a transcriptional reporter for *Spry4*, which serves as a highly sensitive indicator of expression changes linked to signaling-induced transcription during differentiation, I aimed to confirm the involvement of specific signaling systems such as mTOR and FGF. This approach intended to uncover general transcriptional regulators responsible for gene regulation, potentially downstream or independent of the signaling systems.

1.6.3 Understand the role of *Med12* in transcriptional regulation during early differentiation

The genome-wide screen revealed the Mediator subunit *Med12* as a potential regulator of transcription in mESCs, as it strongly promoted the expression of *Spry4*. Focusing specifically on *Med12*, the primary objective was to analyze the relationship between *Med12* and signaling-induced gene regulation during differentiation. Similar to what has been observed with other Mediator subunits, I aimed to determine if *Med12* is essential for the functionality of pluripotency-associated signaling pathways. Furthermore, to understand the functional role of *Med12* during cell type transitions and changes in transcriptional programs, I sought to characterize *Med12*-mutant cells during the exit of pluripotency, primitive endoderm, and anterior visceral endoderm differentiation. Combining the analysis of the effects on the establishment of these cell states involved in symmetry breaking in embryonic development, together with measurements of biological noise levels and plasticity during differentiation, was performed to derive a general function of *Med12*.

2 Materials

2.1 Used oligo sequences

Table 1: Used primer sequences with indicated function.

Primer sequence (5'→3')	Function
CTGACTTGGGTGCTTGAGTGTTTTG	Genotyping Primer for <i>Med12</i> mutants (fwd)
CATTCTCCAGGAAGAGCAGTCAATTC	Genotyping Primer for <i>Med12</i> mutants (rev)
ACTATCCCGCCACCGTTGCCGTC	Genotyping Primer for <i>Fgf4</i> mutant
GATGGAATTCGCACCGAGAG	Genotyping Primer for <i>Fgf4</i> mutant
GGCTAGTCCCTCCTTGCTTCC	Long range PCR Primer for <i>Spry4^{H2B-Venus}</i> line
GGCTGGAGGTCCCTGAACTGC	Long range PCR Primer for <i>Spry4^{H2B-Venus}</i> line
CAAGCAGAAGACGGCATAACGAGATCGGTTCAAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATGCTGGATTGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATTAACGAGTTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATTAACAGTTGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATATACTCAAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATGCTGAGAAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATATTGGAGGGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATTAGTCTAAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATCGGTGACCGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATTACAGAGGGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATATTGCAAGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
CAAGCAGAAGACGGCATAACGAGATTATGTCTTGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT	P7 Indexing in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTCTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTGCTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTAGCTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTCAACTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTTGCACCTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTACGCAACTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen
AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGA TCTGAAGACCCTTGTGGAAAGGACGAAACACCG	P5 Staggering in CRISPR Screen

Table 2: Used gRNAs for indicated genes.

Gene	Name	gRNA Sequence (5'→3')	Sequence for cloning into px459 (5'→3')
Med12	Med12_gRNA_3_Ex7_fw	CAGCTATGACATGAGATGAG	caccgCAGCTATGACATGAGATGAG
Med12	Med12_gRNA_3_Ex7_rev	CAGCTATGACATGAGATGAG	aaacCTCATCTCATGTATAGCTGc
Med12	gRNA_Med12_fw	CATATCTTTACAGTACTCAG	caccgCATATCTTTACAGTACTCAG
Med12	gRNA_Med12_rev	CATATCTTTACAGTACTCAG	aaacCTGAGTACTGTAAGATATGc
Grb2	gRNA_Grb2_fw	AAACACTTACTTGACGGACA	caccgAAACACTTACTTGACGGACA
Grb2	gRNA_Grb2_rev	AAACACTTACTTGACGGACA	aaacTGTCGGTCAAGTAAGTGTTC
Ptpn11	gRNA_Ptpn11_fw	AGAGAACGAAGTCTCCGGGG	caccgAGAGAACGAAGTCTCCGGGG
Ptpn11	gRNA_Ptpn11_rev	AGAGAACGAAGTCTCCGGGG	aaacCCCCGGAGACTTCGTTCTCTc
Sox2	gRNA_Sox2_fw	GATAAGTACACGCTTCCCGG	caccGATAAGTACACGCTTCCCGG
Sox2	gRNA_Sox2_rev	GATAAGTACACGCTTCCCGG	aaacCCGGGAAGCGTGTACTTATC
Tada1	gRNA_Tada1_fw	CCTTCAAGTTGGCTCGCGT	caccgCCTTCAAGTTGGCTCGCGT
Tada1	gRNA_Tada1_rev	CCTTCAAGTTGGCTCGCGT	aaacACGCGAGGCCAAGTGAAGGc
Fam98b	gRNA_Fam98b_fw	CTTGGACAGATCTTACGGG	caccgCTTGGACAGATCTTACGGG
Fam98b	gRNA_Fam98b_rev	CTTGGACAGATCTTACGGG	aaacCCCGTGAAGATCTGTCCAAGc
Med24	gRNA_Med24_fw	TCGACTCCGAGAGGATCACC	caccgTCGACTCCGAGAGGATCACC
Med24	gRNA_Med24_rev	TCGACTCCGAGAGGATCACC	aaacGGTGATCTCTCGGAGTCGAc
Med25	gRNA_Med25_fw	TCTTTGCAGGTACTTCAACG	caccgTCTTTGCAGGTACTTCAACG
Med25	gRNA_Med25_rev	TCTTTGCAGGTACTTCAACG	aaacCGTTGAAGTACCTGCAAAGAc
Eip3	gRNA_Eip3_fw	TACCCAACCCTGGTTATCCG	caccgTACCCAACCCTGGTTATCCG
Eip3	gRNA_Eip3_rev	TACCCAACCCTGGTTATCCG	aaacCGGATAACCAGGGTTGGGTAc
Ikbkap	gRNA_Ikbkap_fw	GATTCATGACCATAACCCCA	caccGATTCATGACCATAACCCCA
Ikbkap	gRNA_Ikbkap_rev	GATTCATGACCATAACCCCA	aaacTGGGGTTATGGTCATGAATC
Eip5	gRNA_Eip5_fw	AGCGAAGAAGAGTTTCGCGA	caccgAGCGAAGAAGAGTTTCGCGA
Eip5	gRNA_Eip5_rev	AGCGAAGAAGAGTTTCGCGA	aaacTCGCGAAACTCTTCTTCGCTc
Kti12	gRNA_Kti12_fw	CGAGGCAGTACAGCTCGTAG	caccgCGAGGCAGTACAGCTCGTAG
Kti12	gRNA_Kti12_rev	CGAGGCAGTACAGCTCGTAG	aaacCTACGAGCTGTACTGCCTCGc
Tsc2	g_Tsc2_fw	TGAACCACATGGCTATGACG	caccgTGAACCACATGGCTATGACG
Tsc2	g_Tsc2_rev	TGAACCACATGGCTATGACG	aaacCGTCATAGCCATGTGGTTCAc
Tsc1	g_Tsc1_fw	ATGGGTACATCCATAAAGG	caccgATGGGTACATCCATAAAGG
Tsc1	g_Tsc1_rev	ATGGGTACATCCATAAAGG	aaacCCTTTATGGGATGTACCCATc
Lamtor4	g_Lamtor4_fw	ATGGTGCAACTGGAAGCCAC	caccgATGGTGCAACTGGAAGCCAC
Lamtor4	g_Lamtor4_rev	ATGGTGCAACTGGAAGCCAC	aaacGTGGCTTCCAGTTGCACCATc
Lztr1	g_Lztr1_fw	CGGATGGCCACACGTAACAG	caccgCGGATGGCCACACGTAACAG
Lztr1	g_Lztr1_rev	CGGATGGCCACACGTAACAG	aaacCTGTTACGTGTGGCCATCCGc
Smarcc1	g_Smarcc1_fw	TGAGCAAGAAGATCTTACCA	caccgTGAGCAAGAAGATCTTACCA
Smarcc1	g_Smarcc1_rev	TGAGCAAGAAGATCTTACCA	aaacTGGTAAGATCTTCTTGCTCAc
Ficn	g_Ficn_fw	GTGTGCAAGAGACTTGAAGC	caccgGTGTGCAAGAGACTTGAAGC
Ficn	g_Ficn_rev	GTGTGCAAGAGACTTGAAGC	aaacGCTTCAAGTCTCTTCGACACc
Med12L	g_Med12L_fw	CACTCGCCCCACATGATAAT	caccgCACTCGCCCCACATGATAAT
Med12L	g_Med12L_rev	CACTCGCCCCACATGATAAT	aaacATTATCATGTGGGGCGAGTGc

2.2 Commercial kits

Kit	Function
3' CellPlex Kit Set A (10x Genomics, #1000261)	Cell labeling for multiplexed single-cell RNA sequencing
3' Feature Barcode Kit (10x Genomics, #1000262)	Sample preparation for multiplexed single-cell RNA sequencing
Chromium Next GEM Chip G Single Cell Kit (10x Genomics, #1000127)	Sample preparation for multiplexed and non-multiplexed single-cell RNA sequencing
Dual Index Kit TT Set A (10x Genomics, #1000215)	Library preparation for multiplexed single-cell RNA sequencing
Dual Index Kit NN Set A (10x Genomics, #1000243)	Library preparation for multiplexed single-cell RNA sequencing
Monarch Plasmid Miniprep Kit (NEB, # T1010S)	Purification of small amounts of plasmid DNA
QIAGEN Plasmid Midi Kit (Qiagen, # 12143)	Purification of larger amounts of plasmid DNA
DNA- free™ DNA Removal Kit (Invitrogen, #AM1906)	DNA removal from RNA samples
Terra™ PCR Direct Genotyping Kit (Takara, #639285)	DNA purification and PCR amplification for genotyping
Monarch® Genomic DNA Purification Kit (NEB, #T3010S)	Genomic DNA purification
Single Index Kit T Set A (10x Genomics #1000213)	Sample preparation for non-multiplexed single-cell RNA sequencing
Chromium Next GEM Single Cell 3' Kit v3.1 (10x Genomics, #1000269)	Sample preparation for multiplexed single-cell RNA sequencing
Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1 (10x Genomics, #1000128)	Sample preparation for non-multiplexed single-cell RNA sequencing

3 Methods

3.1 Cell culture

Routine culture of mESCs was performed at 37 °C with 5% CO₂. Media conditions varied between cell lines and experiments. In general, for routine culture, mESCs were kept in either serum + LIF medium (ES + LIF), composed of GMEM with 10% fetal bovine serum (FBS), 2 mM GlutaMAX, 1 mM sodium pyruvate, 0.1 mM β-mercaptoethanol and 10 ng/mL LIF), or in 2i + LIF. 2i + LIF was N2B27 supplemented with 1 μM PD0325901 (SelleckChem), 3 μM CHIR99021 (Tocris), and 10 ng/ml LIF (prepared in-house). N2B27 was prepared as a 1:1 mixture of DMEM/F12 and Neuropan Basal Medium (both from PAN Biotech), supplemented with 0.5X N2 and 0.5X B27 supplements, 1X L-Glutamax, 0.0025% BSA, and 0.1 mM β-mercaptoethanol (all from ThermoFisher). *Fgf4*-mutant cell lines were cultured in 2i + LIF supplemented with 10% FBS. Cells were maintained on plastic flasks and dishes (Sarstedt) coated with 0.1% gelatine when cultured in ES + LIF medium or with fibronectin in all other media for > 20 min. Passaging every 2 to 3 days was performed by singularizing cells with trypsin (PAN Biotech) or Accutase (Sigma-Aldrich). Media was exchanged daily.

3.2 Cell lines

All cell lines generated in this study were derived from the E14tg2a wild-type line (Hooper et al., 1987). The GATA4-mCherry inducible line used for single-cell RNA sequencing has been described previously (Raina et al., 2021). The *Spry4*^{H2B-Venus/+}-reporter line was generated with a previously described targeting construct (Morgani et al., 2018) using lipofectamine 2000 according to the manufacturer's instructions (Thermo Fisher Scientific). Briefly, 1 μg vector DNA was added to 50 μl Opti-Mem (Gibco), and 4 μl Lipofectamine 2000 were added to another 46 μl Opti-Mem. Both mixtures were combined and vortexed for 30 s, incubated for 10 min, and used to resuspend 5 * 10⁵ singularized cells of the parental line. After 10 min incubation, cells were seeded at clonal density and kept in KH2 medium (KO-DMEM with 1x non-essential amino acids, 2 mM Glutamax, 0.1 mM β-mercaptoethanol (all Thermo Fisher), 15% FBS and 10 ng/ml LIF) on mitotically inactivated feeder cells. 24 h after transfection, neomycin selection was started. Clones were picked 5 to 6 d after transfection and expanded. Correctly targeted clones were identified via long-range PCR as described in Morgani et al., 2018 and further expanded without the presence of feeder cells. GATA6-mCherry inducible lines were established as described for GATA4-mCherry inducible lines in (Raina et al., 2021), but replacing the *Gata4* with a *Gata6* coding sequence in the PiggyBac vector for inducible gene expression. Transfection and the establishment of multiple

clonal lines were performed similarly to all further cell line generation without the presence of feeder cells. Multiple Clones were tested for GATA6-mCherry induction levels upon dox-treatment by flow cytometry. Three independent clones with induction levels similar to or slightly higher than the previously established GATA4-mCherry inducible lines were selected for the experiment shown in Figure 22D, and a single clonal line was chosen for all other experiments. Newly generated *Spry4^{H2B-Venus/+}*-reporter and GATA6-mCherry inducible cell lines were checked for karyotypic abnormalities. To label nuclei for time-lapse imaging, cells were transfected with pCX-H2B-Cerulean-IRES-puro (Schumacher et al., 2023). Cell lines carrying PiggyBac transgenes were kept under constant appropriate selection to prevent transgene loss over passaging.

3.3 sgRNA cloning and generation of single-gene mutant mESCs

For mutagenesis of individual genes via CRISPR/Cas9, gene-targeting sgRNAs (Table 2) were cloned into pX459 (Addgene plasmid #48139) using BbsI (NEB) overhangs following Ran et al., 2013 (Ran et al., 2013). Briefly, the top and bottom strands of gRNA oligos were first annealed and phosphorylated using T4 PNK (NEB). The *Staphylococcus aureus* Cas9 (SpCas9)-containing vector was cut with BbsI, and the oligo duplex was ligated into the backbone with T7 ligase (NEB) in a cyclizing reaction. Exonuclease treatment digested residual linearized vector before transformation into competent XL10-gold bacteria. Ampicillin-resistant bacteria colonies were used to inoculate overnight cultures, plasmid DNA was purified subsequently, and insertion of the gRNA was verified by Sanger sequencing. Clonal mutant lines were generated using a combination of sgRNAs with targeting sequences 100 to 200 bp apart in the genome. Single sgRNAs were used when generating polyclonal lines. For validation experiments of the CRISPR screen (Figure 12 and 13), I selected the most enriched sgRNA in sorted cells. According to the manufacturer's protocol, a total of 1 µg of sgRNA containing pX459 vectors was mixed with a final concentration of 0.04 µg/ml Lipofectamine 2000 (Thermo Fisher Scientific) in Opti-MEM (Gibco). For the generation of clonal lines, cells were seeded at clonal density into 10 cm dishes after transfection; for polyclonal experiments, approximately 5×10^4 cells/cm² were seeded. To enrich for successfully transfected cells, selection with 1.5 µg/ml puromycin was started 24 h after transfection for 48 h. To establish clonal lines, single-cell derived colonies were picked 4 to 6 d after transfection and expanded. Genomic DNA was purified for molecular characterization of genetic lesions with Terra™ PCR Direct Genotyping Kit (Takara), followed by PCR amplification (Table 1) and Sanger sequencing of specific genomic regions encompassing the target site.

3.4 Genome-wide CRISPR Screen

To generate stably CAS9-expressing *Spry4*^{H2B-Venus/+} reporter cells, cells were transduced with lentiCas9-Blast lentiviral particles (Addgene #52962-LV) at a multiplicity of infection of approximately 0.1. Transduction was performed with attached cells 20 h after seeding at 2×10^4 cells/cm² in the presence of 5 µg/ml Polybrene in ES + LIF. Continuous blasticidin (15 µg/ml, Gibco) selection was started 24 h after transduction. Lentiviral particles of the genome-wide gRNA library Brie (Addgene #73633) were generated according to standard protocols (Doench et al., 2016). For library transduction, 150×10^6 CAS9-expressing *Spry4*^{H2B-Venus/+} reporter cells were detached and mixed with the virus library in ES + LIF with 5 µg/ml Polybrene. The following day, the same number of cells was reseeded and put under selection with puromycin (1.5 µg/ml, Sigma-Aldrich). Comparing cell counts with and without selection indicated a transduction efficiency of 25%, resulting in a > 400-fold coverage of transduced cells per gRNA. At least 31×10^6 cells were processed in subsequent steps to maintain gRNA coverage.

To identify gRNAs enriched in cell populations with high and low *Spry4*:H2B-Venus expression, at least 0.5×10^6 cells with the lowest or highest 1% of *Spry4*:H2B-Venus fluorescence, or 3×10^6 cells with the lowest or highest 5% of *Spry4*:H2B-Venus fluorescence were FAC sorted and their DNA isolated by column-based genomic DNA purification (Monarch Genomic DNA Purification Kit, NEB). For reference, the genomic DNA of 31×10^6 non-sorted control cells was purified in parallel. The integrated gRNA was PCR amplified using Pfu polymerase (prepared in-house) with a sample-specific sequencing adapter and index-containing primers (Table 1; Carlini et al., 2021) using the complete purified genomic DNA as the template. PCR samples were purified with the SPRIselect reagent (Beckman Coulter) with double-sided size selection. Briefly, 0.5x SPRIselect was added to each sample, incubated for 5 min at RT, and the SPRIselect was removed with a magnet. This supernatant was again mixed with 1.2x SPRIselect, incubated, and then discarded. After washing the beads, the DNA library was eluted from the beads and used for sequencing.

Paired-end Illumina Sequencing with a read length of 150 bp pairs was performed with at least 10×10^6 reads per sorted sample and 30×10^6 reads for the unsorted library controls. The raw reads were trimmed using Cutadapt (Martin, 2011) to remove the vector binding sequence. The reads were mapped to individual gRNAs, counted using *norm-method total*, and statistically tested on the targeted gene levels using *gene-lfc-method alphamean* with Mageck (W. Li et al., 2014). Hits were selected based on the false discovery rate.

3.5 Immunostaining

Immunostaining was performed as previously described (Schröter et al., 2015). Briefly, cells were washed with PBS containing Calcium and Magnesium, followed by fixation with 4% paraformaldehyde (Histofix, Sigma-Aldrich) for 15 min. Cells were permeabilized and blocked by rinsing and washing three times with PBS with 0.1% Triton X-100 and 1.0% bovine serum albumin (PBT-BSA). Primary antibodies (anti-mouse NANOG (Affymetrix eBioscience, Cat.:14-5761), anti-SOX17 (R&D systems, Cat.: AF1924), anti-GATA6 (R&D systems, Cat.: AF1700), anti-Oct3/4 (POU5F1, Santa Cruz Biotechnology, Cat.: sc-5279, 1:100), anti-EOMES (abcam, Cat.: ab23345), anti-OTX2 (Neuromics, Cat.: GT15095) were diluted 1:200 if not indicated otherwise in PBT-BSA and incubated with the cells overnight at 4 °C. The next day, cells were washed in PBT-BSA and incubated with Alexa Fluor-conjugated secondary antibodies at 4 µg/ml (Invitrogen/Life Technologies) and Hoechst 33342 at 1 µg/ml (Invitrogen) in PBT-BSA in the dark for 2 h. Finally, samples were rinsed and washed with PBS and imaged in a mounting medium consisting of 80% glycerol, 16% PBS, and 4% n-propyl-gallate.

Similarly, immunostaining for FACS analysis was performed on cells detached from the culture dishes with Accutase. PBS + 1% BSA + 0.25% Saponin was used for permeabilization (30 min at RT), washing (3x after each antibody incubation), and staining (with antibodies from above). Cells were spun down at 200 xg for 5 min after each wash. Before analysis at the cytometer, cells were passed through a cell strainer (FALCON, 35µm mesh).

3.6 Immunoblotting

For western blot analysis of MED12 and phospho-ERK (ppERK), cells were washed twice with ice-cold with 1 mM activated orthovanadate in case of ppERK detection. Cells were mechanically detached in lysis buffer, based on commercially available lysis buffer (Cell Signaling) supplemented with benzonase (Sigma-Aldrich), cOmplete EDTA-free protease inhibitor cocktail (Roche), phosphate inhibitors P1 and P2 (Sigma). The lysates were snap-frozen in liquid nitrogen twice and centrifuged. Protein concentration in the supernatant was measured with a micro BCA assay (Thermo Scientific). For western analysis, 20 µg of protein per sample were denatured by adding 5x Laemelli buffer and incubation at 95 °C for 5 min. The SDS-PAGE was run in 1x MOPS buffer (ThermoFisher) with 5 mM sodium-bisulfate and immediately transferred onto methanol-activated PVDF membranes. Transfer was performed in transfer buffer (12mM Tris-Base, 96mM Glycine, 20% methanol) at 40 V for 1.5 h in a NuPage transfer system (ThermoFisher). Membranes were blocked at RT for 1 h in Intercept blocking buffer (LI-COR), which was also used for the dilution and incubation with the primary antibodies anti-Tubulin 1:10000 (T6074, Sigma), anti-

ppERK1/2 1:1000 (4370S, Cell Signaling), anti-total ERK1/2 1:1000 (ab36991, Abcam) or anti-MED12 (1:250, A300-774A, Bethyl Laboratories). Appropriate secondary antibodies (IRDyes, LI-COR) were used at a dilution of 1:5000. Blots were imaged on an Odyssey CLx (LI-COR). Quantification of ppERK bands was performed using the gel quantification tool in Fiji, using the combined intensities of ERK1 and ERK2 bands, and normalizing ppERK by total ERK signals.

3.7 In situ HCR

In situ hybridization probes against the spliced mRNAs of *Spry4* and *Nanog* were designed by, and all commercial reagents for the staining procedure were obtained from Molecular Instruments, if not indicated otherwise. mRNAs were stained in solution as previously reported (Choi et al., 2018). Specifically, cells from confluent 6-well plate wells were washed with PBS and detached using trypsin. Following centrifugation for 5 min at 200xg, the specification for all further centrifugation steps, cells were fixed with 4% PFA for 1 h. After removal of the fixative via centrifugation, cells were washed four times with PBS with 0.1% Tween20 (PBST, Bio-Rad). Cells were incubated overnight in ice-cold ethanol. Starting with 1×10^6 cells, and two additional washes with PBST, cells were incubated in 400 μ l probe hybridization buffer for 30 min at 37 °C. 2 pmol of gene-specific hairpins were prepared in 100 μ l preheated probe hybridization buffer and added to the cell suspension, which was subsequently incubated overnight at 37 °C. Preheated wash buffer was used to remove residual hybridization oligos by washing thrice with 10 min of intermediate incubation. One final wash and 5 min incubation were performed in saline-sodium citrate (SSC, Sigma-Aldrich) buffer with 0.1% Tween20. Cells were resuspended in 150 μ l amplification buffer. Spectrally compatible labeled hairpins were formed by mixing 10 pmol of both corresponding hairpins, heating them to 95 °C for 90 s, and cooling them back to RT. Hairpins were added into 100 μ l of amplification buffer, mixed with the cell suspension and incubated in the dark for 1 h. A final five washes with SCCT were followed by mounting the sample onto a microscopy glass slide. Samples were resuspended in 20 to 100 μ l SCCT to ensure high single-cell densities. Only 2 μ l were squished with high manually applied force between the slide and a cover slip to maximize area per cell and distance between mRNA spots.

3.8 Fluorescence microscopy and image analysis

3.8.1 Image acquisition

Tilescans of immunostainings were imaged with a Leica SP8 confocal microscope (Leica Microsystems) with a 63x 1.4 NA oil immersion objective. Images of live *Spry4*^{H2B-Venus/+}-reporter

cells were acquired with an Olympus IX81 widefield microscope equipped with a stage top incubator (ibidi), pE4000 illumination (CoolLED), ORCA-Quest qCMOS camera (Hamamatsu) with a 63x 1.35 NA oil immersion objective. Hardware was controlled by Olympus CellSens Software. Images of in situ HCR for mRNA counting was acquired on the same system with a 100x 1.4 NA objective with 1024 pixels in X- and Y-dimension and at least 20 z-slices, separated by 0.4 μm . Only well-separated single cells were selected to circumvent the need for segmentation during downstream processing.

Time-lapse imaging was performed with a 40x 0.9 NA objective on an Olympus IX81 widefield microscope equipped with an LED-based illumination system (pE4000, CoolLED) and an iXon 888 EM-CCD camera (Andor). MicroManager (Edelstein et al., 2010) was used to control the hardware. Images were taken every 10 min.

3.8.2 Image analysis

In general, images were analyzed in Fiji (Schindelin et al., 2012). For quantification of immunostainings, segmentation was performed with StarDist 2D (Schmidt et al., 2018) using the versatile (fluorescent nuclei) model and default post-processing parameters. Mean fluorescence intensity was measured in segmented cells in all acquired channels. Cells with a nuclear area smaller than 40 μm^2 were filtered out. To determine fluorescence intensity threshold values for the classification of cell types, I manually selected thresholds that best bisected the bimodal expression profiles of the lineage markers. The same thresholds were applied to different samples in a single experiment.

Counting mRNA spots from in situ HCR was performed similarly to previously established protocols (Raj et al., 2008). Spots were detected on z-projections in 2D. First, to segment single cells, a median z-projection was blurred using a Gaussian blur ($\sigma = 3$), allowing robust manual thresholding of the cells. “Analyze Particles” was used to detect cells completely localized in the field of view within a $2 * 10^4$ to $15 * 10^4$ pixel size range and a minimum circularity of 0.5. Within each cell for every channel, spots were enhanced from a maximum intensity projection with a Top Hat and Mexican Hat filter. The *Find Maxima* function detected spots with iterating prominence values. The detected number of spots was plotted against prominence values for each cell, and the highest prominence value falling onto an intermediate plateau of spot counts in most cells was selected for all images of one replicate and channel. Spot counts per cell were exported from Fiji and plotted in R as cumulative distributions.

Tracking was performed with the manual tracking function in Trackmate v7 (Ershov et al., 2022), and fluorescence intensity was measured as the mean intensity in a spot with a 4 μm radius within

the nucleus. Cells tracks were manually linked to cells in the immunostaining based on their location and shape at the last acquired time frame. For analysis of the induction levels per cell, cells from every second frame were segmented using StarDist 2D (Schmidt et al., 2018) based on the H2B-Cerulean signal and mCherry intensity measured within Nuclei. R, within RStudio, was used for downstream analysis of tracks, quantifications, and plotting. Tracks were smoothed with a rolling average over 7 frames. For ROC analysis, the R package pROC (Robin et al., 2011) was applied for each time point, and the optimal threshold was defined by Youden's J statistic (Youden, 1950) for each time point.

3.9 Flow cytometry

Adherent cells analyzed or sorted live, were washed with PBS and detached with trypsin or Accutase. After centrifugation (5 min at 200xg), cells were resuspended in PBS supplemented with 0.5% BSA. Before analysis at the cytometer, cells were passed through a cell strainer (FALCON, 35 μ m mesh). Analysis of *Spry4*:H2B-Venus reporter expression in live or fixed cells and fixed immunostained cells was performed on an LSRII flow cytometer (BD Biosciences). Cell sorting and analysis of GATA6-mCherry expression was carried out using a FACS Aria Fusion (BD Biosciences). Primary data analysis, including gating single cells based on side scatter (SSC) and forward scatter (FSC), was done with FlowJo version 9 (BD Biosciences).

3.10 Clonogenicity Assay

Clonogenicity assays were performed according to Kalkan et al., 2017. Briefly, 1×10^4 cells/cm² were seeded in 2i + LIF for 24 h, followed by differentiation in N2B27 for 48 h. Control wells for each parental cell line were kept in 2i + LIF. Cells were then detached with Accutase to single cells, and 500 cells were reseeded into 6-well plates with 2i + LIF + 10% FBS. 10% FBS were included for all cell lines to support survival of *Fgf4*-mutant cells and maintain comparability among conditions.

After 5 d, the colonies formed were fixed and stained for alkaline phosphatase (Sigma-Aldrich) to distinguish pluripotent and differentiated colonies. Tile scans of the wells were acquired with an Olympus IX81 widefield microscope with a 4x 0.16 NA objective. I applied background subtraction, gaussian blurring, Otsu-thresholding, and conversion of images to a binary mask in ImageJ, and then used the *AnalyzeParticles* function to set thresholds for size and circularity and to determine the number of colonies. Colony numbers were normalized to the number of colonies obtained in the control.

3.11 Bulk RNA sequencing

3.11.1 Experimental implementation

Three different bulk RNA sequencing experiments were performed in this study. First, E14 *Fgf4*-mutant cells were seeded at 45×10^3 cells/cm², grown in N2B27 + Chiron + LIF for 18 h, and stimulated with 0, 0.625, 1.25, 2.5, 5, 10, 20 or 40 ng/ml FGF4 in N2B27 + 1 µg/ml heparin for 6 h. Second, *Spry4*^{H2B-Venus/+}-reporter cells and *Med12*-mutant cells of this line were seeded at 35×10^3 cells/cm², grown in 2i for 14 h, and differentiated in N2B27 or kept in 2i as controls for 24 h. Triplicates of these experiments were obtained by seeding different passages on different days. Finally, 55×10^3 cells/cm² of *Fgf4*-mutant cells and *Fgf4 Med12* double mutant cells were seeded into N2B27 + Chiron + LIF and differentiated with or without 40 ng/ml FGF4 in N2B27 + 1 µg/ml heparin for 6 h. Here, triplicates of different passages were complemented by triplicates of different clonal lines of the *Med12* mutant during the preparation of the first replicate.

RNA isolation was performed with TRIzol (ambion) according to the manufacturer's instructions. Briefly, RNA was extracted by homogenizing cells in 1 ml Trizol at RT for 5 minutes. 0.2 ml Chloroform was added, and the mixture was shaken, followed by incubation at RT for 2 to 3 minutes. Centrifugation at 12,000 g for 15 minutes at 4 °C separated the lysate into organic and aqueous phases. From the aqueous phase, RNA was precipitated by adding 0.5 ml isopropyl alcohol and incubating at RT for 10 minutes. After centrifugation at 12,000 g for 10 minutes at 4 °C, the RNA pellet was washed with 1.2 ml 75% ethanol. The ethanol was removed after vortexing and centrifugation at 7,500 g for 5 min at 4 °C. The isolated RNA was dissolved in 40 µl RNase-free water and incubated for 10 minutes at 55 to 60 °C. From total RNA, sequencing libraries were prepared with polyA-enrichment, followed by paired-end sequencing at a read-length of 150 bp and depth of approximately 30×10^6 reads per sample. Strand-specific libraries were generated only for the FGF-titration experiment (Figure 6) and the differentiation of the *Med12* wild-type and mutant cells (Figure 15).

3.11.2 Bulk RNA sequencing data analysis

Raw reads were mapped to the mouse genome (GRCm39, release 108 (both *Med12* mutant experiments) or release 97 (FGF titration experiment) with hisat2 (v2.1.0; Kim et al., 2019). SeqMonk was used to quantify counts per gene, either as transcript per kilobase million (TPM) or as raw counts as input for downstream DESeq2 analysis (Love et al., 2014) to identify differentially expressed genes. For the definition of FGF target genes from the FGF-titration experiment, since this experiment was not performed with multiple replicates, the variation in gene expression compared to the change upon FGF signaling was estimated using the different

FGF concentrations. In detail, the gene list was first limited to genes with an expression within the 70% to 100% most expressed genes in at least one sample to minimize the influence of technical noise of lowly expressed genes. Second, a threshold for a minimal 2-fold change was applied. Finally, a logistic regression (using the R function `LL.4`) of the expression values of these genes had to converge to ensure an FGF response was not dominated by technical noise. The concentration at which the fit reached half-maximal expression, was used to determine the FGF4 sensitivity of single genes.

Gene Ontology (GO)-term enrichment analysis was performed with ShinyGO 0.77 (Ge et al., 2020). All statistically significant differentially expressed genes from the indicated conditions, irrespective of their direction of regulation, were used to find enriched biological processes annotated in the GO database. Biological processes were thresholded for a minimal sorted by their false discovery rate (FDR).

Differentiation delay in *Med12* mutants was estimated according to (Lackner et al., 2021). I first determined the expression change of the naïve marker genes *Nanog*, *Esrrb*, *Tbx3*, *Tfcp2l1*, *Klf4*, *Prdm14*, and *Zfp4* in *Med12*-mutant and wild-type cells and then plotted the Euclidean distance of this expression change to that of the time-resolved dataset from Lackner et al., 2021.

Signaling footprint analysis in *Med12* mutants was performed similarly to Lackner et al., 2021. This study defined a specific set of target genes for each pluripotency-associated signaling system based on gene expression changes in knockouts of signaling genes. A signaling footprint for a knockout line can then be determined from the difference in the expression of pathway footprint genes to the wild-type line after 24 h of differentiation. Measures for the signaling footprint are the Spearman correlation between each knockout line and the respective pathway defining knockout and the ratio between the sum of expression fold changes between a knockout line and the respective pathway defining knockout, defined as pathway activity. To compare the *Med12* mutant data from this study, the wild-type conditions were used for batch correction.

3.12 Single-cell RNA sequencing experimental procedure and analysis

3.12.1 Droplet-based single-cell RNA sequencing and burst parameter estimation

To analyze the transcriptional dynamics upon FGF4 stimulation with single-cell RNA sequencing, 1.5×10^4 *Fgf4*-mutant cells/cm² were seeded in 6-well plates in ES + LIF medium. The next day, 24 h of FGF stimulation with 0, 5, or 50 ng/mL FGF4 was started in N2B27 medium with 1 µg/ml Heparin and Chiron. For GEM generation, cells were washed, detached, and strained as during preparation for FACS analysis (see 3.9).

Droplet generation, lysis, mRNA and cell barcode capture, and generation of the gene expression libraries was performed following the instructions by 10x genomics (Chromium Next GEM Single Cell 3' Reagent Kits v3.1 with Feature Barcoding technology for CRISPR Screening, CG000205 Rev D). Specifically, I used 4.1 μl of a cell suspension with 400 cells/ μl for a targeted cell recovery of 1000 cells per sample. I chose 12 PCR cycles for cDNA amplification and 13 cycles for the sample index PCR. Concentration and insert size distribution for both the gene expression library and the cell multiplexing library were determined with a BioAnalyzer High Sensitivity DNA Assay (Agilent). Sequencing was performed on a NovaSeq 6000 on multiple flowcells with a paired-end 150 bp configuration. Between $3.5 * 10^8$ and $1.5 * 10^8$ read pairs were obtained for each gene expression library.

Alignment of the raw reads to the mouse genome mm10 (GENCODE vM23/Ensembl 98, obtained from 10x Genomics) was performed with CellRanger (version 4.0.0, 10x Genomics). Downstream analysis was performed in R with Seurat v5 (Hao et al., 2023). I first filtered cells by removing barcodes with ≤ 4000 detected genes and $\geq 10\%$ of reads aligned to mitochondrial genes, retaining between 361 and 808 cells per sample with median mRNA counts per cell between 93557 and 109942 in the different samples (Supplementary Table 1). mRNA counts for each gene were normalized using the Seurat function *SCTransform*. Briefly, the detected counts were normalized by performing a regularized negative binomial regression to stabilize variance across expression levels while maintaining biological heterogeneity. FGF target genes were defined with the FindMarker function, a minimal log₂-fold change of 0.5, and an adjusted p-value < 0.05 . As control gene sets, the same number of genes with the most similar expression levels (in both directions) to the defined FGF target genes were selected. A minimal average detection of 2.5 normalized counts per cell was chosen as a minimal requirement for burst parameter estimation with *txburst* (Larsson et al., 2019). Maximum likelihood point estimates were inferred for k_{on} (burst frequency, in units of the degradation rate) and $k_{\text{syn}}/k_{\text{off}}$, reflecting the burst size and filtered for a minimal fit quality as in Larsson et al., 2019. The coefficient of variation was calculated for every gene of the same gene sets as the standard deviation of all detected counts per gene divided by its mean.

3.12.2 Multiplexed single-cell RNA sequencing

In scRNAseq to compare PrE differentiation in wild-type and *Med12*-mutant cells, cells were seeded at a $3.5 * 10^4$ cells/ cm^2 density in 6-well plates in 2i + LIF and grown overnight. The following day, induction in 2i + LIF + dox was first started in the mutant clones and 4 h later in the wild-type lines. After 8 h and 4 h, respectively, Induction was stopped by washing once with N2B27, followed by 20 h of differentiation in N2B27. Controls for each cell line were continuously

kept in $2i + LIF$. For sequencing, cells were washed three times with PBS and detached with Accutase. Accutase was removed by centrifugation, and $1 * 10^6$ cells per sample were resuspended in PBS + 0.04 % BSA and immediately used for multiplexing labeling following the protocol of 10x Genomics for samples with a viability above 80 % (Cell Multiplexing Oligo Labeling for Single Cell RNA Sequencing Protocols with Feature Barcode technology, CG000391). Briefly, cells were spun down once again, resuspended with individual cell multiplexing oligos (CMO no. 301 to 310), and incubated for 5 min at RT. Cells were washed twice with PBS + 1 % BSA and passed through a cell strainer (FALCON, mesh size 35 μm). A total of $1.2 * 10^5$ single cells from all samples were pooled at equal ratios, and $4 * 10^4$ were used for droplet generation, corresponding to a target number of $2.4 * 10^4$ cell-containing droplets. Droplet generation, lysis, mRNA and cell barcode capture, and generation of both the gene expression library as well as the cell multiplexing library were performed following the instructions by 10x genomics (Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index) with Feature Barcode technology for Cell Multiplexing, CG000388). Specifically, I chose 11 PCR cycles for cDNA amplification and 10 cycles for the sample index PCR. Concentration and insert size distribution for both the gene expression library and the cell multiplexing library were determined with a BioAnalyzer High Sensitivity DNA Assay (Agilent). Sequencing was performed on a NovaSeq 6000 on multiple flowcells with a paired-end 150 bp configuration. In total, $1.2 * 10^9$ and $2.3 * 10^8$ read pairs were obtained for the gene expression and multiplexing library, respectively.

Demultiplexing to the individual samples, based on the cell multiplexing barcode and alignment to the mouse genome mm10 (GENCODE vM23/Ensembl 98, obtained from 10x Genomics), was performed with CellRanger (version 7.1.0, 10x Genomics). Downstream analysis was performed in R with Seurat v5 (Hao et al., 2023). I first filtered cells by removing barcodes with ≤ 2500 detected genes and ≥ 15 % of reads aligned to mitochondrial genes, retaining between 1100 and 1700 cells per sample with median mRNA counts per cell between 23233 and 27890 in the different samples (Supplementary Table 3). mRNA counts for each gene were normalized by dividing its counts by the total number of counts per cell, multiplied by 10000. Log1p transformation was applied before plotting expression data as violin plots. For downstream analysis and representation of gene expression as heatmaps, centering counts for each feature and scaling to its standard deviation were applied. Principal component analysis was performed on the 2000 most variable features in the relevant subset of cells. The resolution of the Louvain clustering algorithm was set to 0.05 when clustering multiple samples. In the case of Jaccard-Index estimation, the clustering resolution was set to 0.15, and the clustering was performed on each sample separately. 100-fold repetition of this clustering approach with a random subset of the data with 70 % of the cells allowed the calculation of a Jaccard index, as previously described (Tang et al., 2021). For annotation of the Epi- and PrE-fate, the cells of the differentiated samples

were integrated with Seurat Integration based on the *rpca* reduction. Differentially expressed genes between cell states and genotypes were identified with the *FindMarkers* function in Seurat with a minimal expression difference in the log_{1p} transformed expression values of 0.5. Cell cycle scoring (S- and G2M-Score) was performed using the *CellCycleScoring* function in Seurat. As marker genes for S and G2M phase, mouse homologs of previously published human datasets were used (Schwabe et al., 2020). Biological noise was quantitated and distinguished from modeled technical noise in local neighborhoods of each cell with VarID2 (Rosales-Alvarez et al., 2023).

3.13 Data and code availability

The sequencing data included in this thesis was also partly included in the preprint on bioRxiv and has been deposited at GEO with accession number GSE253609. The single-cell RNA-sequencing dataset of *Fgf4*-mutant cells stimulated with different concentrations of FGF4 has been deposited at GEO with accession number GSE256059. Examples of ImageJ macros used for the main imaging analysis are available on github.com/mfernkorn/ImageJ_Macro_Collection.

4 Results

4.1 Transcriptional responses to FGF signaling in mESCs

4.1.1 FGF signaling induces strong transcriptional changes

During development, extracellular signals trigger gene expression changes that make cells transit between identities. One signal triggering different fates is FGF, which is essential for the differentiation of Epi and PrE cells from ICM cells. In the presence of the transcription factor GATA, higher FGF signaling levels promote the PrE identity, while lower FGF signaling levels are needed to establish Epi cells (Raina et al., 2021; Schröter et al., 2015). Therefore, I asked if specific classes responding to different FGF concentrations explain the formation of the two fates. In pluripotent mESCs, FGF4 is the only expressed FGF and functions in an auto- and paracrine manner. Mutation of the *Fgf4* locus allows the precise control of signaling levels provided as supplemented FGF4. To identify how this affects the transcriptome globally, I performed a bulk RNA sequencing experiment in the *Fgf4*-mutant line comparing different supplemented concentrations of FGF4 after 6 h stimulation (Figure 6A), minimizing the contributions of secondarily induced changes in transcription. The supplemented FGF4 concentrations entirely covered the dynamic range as well as the saturation regime in previous experiments (Raina et al., 2021).

Comparing the transcriptomes of unstimulated and stimulated cells, 588 different protein-coding genes were significantly differentially expressed (Figure 6B). While there was a wide range of sensitivities for upregulated genes upon FGF4 stimulation, downregulated genes were not further downregulated with FGF concentrations above 5 ng/ml (Figure 6B). Single genes showed fold changes between no FGF4 and the highest FGF4 concentration of almost 900-fold for up- (Figure 6C) and 70-fold for downregulated genes (Figure 6D). The effect sizes between down- and upregulation cannot directly be compared since the fold change does not exclusively reflect true underlying biological processes but also depends on the limited sequencing depth for lowly expressed genes. This might result in smaller expression changes for downregulated genes. The smooth increase along increasing FGF concentration, especially for the upregulated genes, not only emphasized the robustness of the RNA sequencing approach but also showed how well populations of cells can discriminate different concentrations of extracellular signals. Among the upregulated genes, there were previously described target genes for FGF signaling, for example, *Etv4*, *Dusp6*, and *Spry4* (Morgani et al., 2018), with *Spry4* being the strongest upregulated gene in response to FGF4 (Figure 6C). Together, this shows that stimulation with FGF4 goes along with major transcriptional adaptations, both in the number of changed genes and their individual fold changes. However, the identified target genes do not classify according to their sensitivity towards

FGF4, which argues against different sensitivity as the explanation for FGF4 concentration-dependent differentiation outcome.

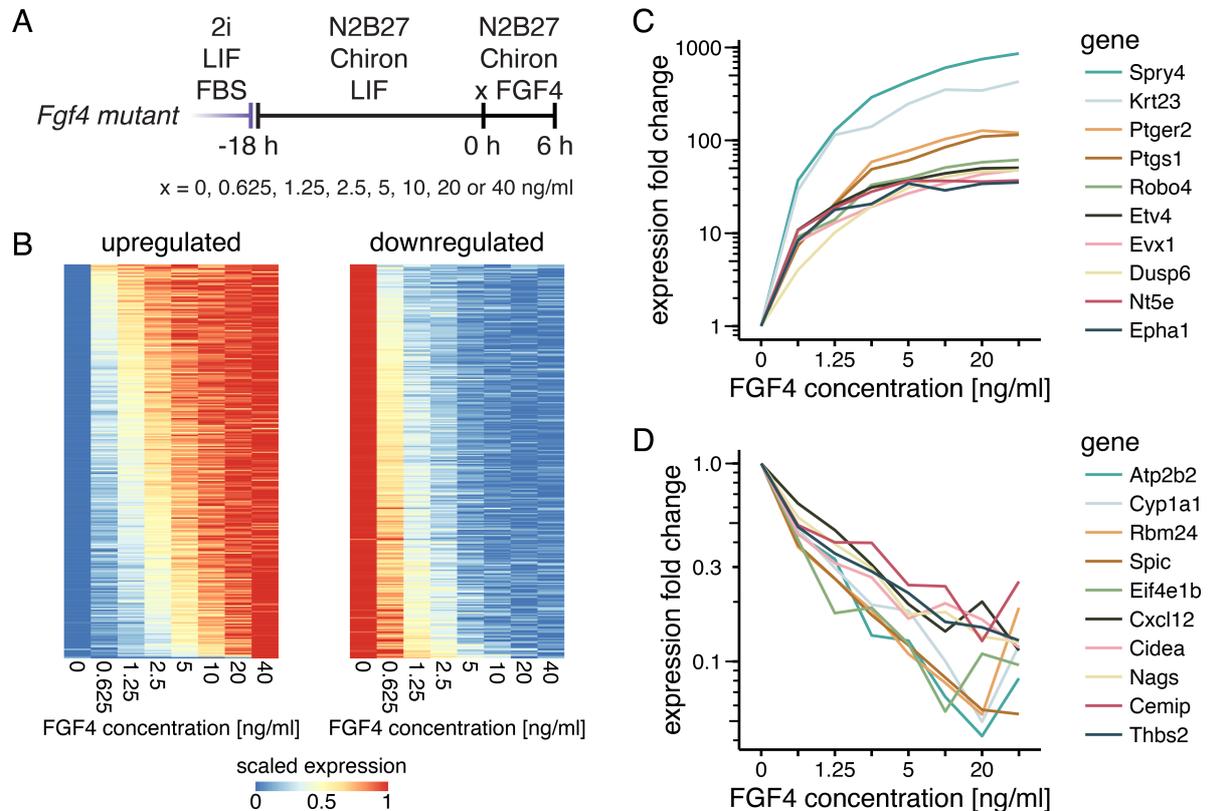


Figure 6: FGF4 stimulation induces strong gene expression in mESCs.

A Schematic of experiment to identify *Fgf4*-regulated genes by bulk RNA sequencing. *Fgf4*-mutant cells were transitioned from 2i + LIF medium containing 10% FBS to N2B27 supplemented with Chiron and LIF for 18 h, followed by stimulation for 6 h with indicated concentrations of FGF4 in N2B27 with Chiron.

B Scaled gene expression (between 0 and 1 for each gene) of differentially expressed genes upon FGF4 stimulation ordered by sensitivity (see chapter 3.11.2). 282 upregulated genes (fold change between 0 and 40 ng/ml FGF4 > 2, left) and 306 downregulated genes (fold change < 0.5, right) with a coherent FGF4 response were detected.

C, D Expression fold change of the ten most upregulated (C) and downregulated (D) genes upon FGF4 titration in *Fgf4* mutants.

4.1.2 Transcriptional dynamics of FGF target genes

The bulk RNA sequencing screen revealed the individual concentration dependencies for all FGF4 target genes in the context of mESCs exiting from pluripotency. Next, I wanted to find out how homogenous the response is at the single-cell level. Potential heterogeneity might explain the graded responses observed on a single gene level. I used droplet-based single-cell RNA sequencing to investigate the FGF-dependent transcriptional response with single-cell resolution upon 24 h of stimulation with FGF4 in an *Fgf4*-mutant mESC line (Figure 7A). High sequencing coverage was ensured for high sensitivity for heterogeneous expression of even lowly expressed

genes, detecting 1882 cells across all three conditions and about 9.3 to 10.1×10^5 median mRNA molecules per cell (Supplementary Figure 6, Supplementary Table 1). This number is equivalent to one-ninth to one-fifth of the total mRNAs in mouse stem cells (Carter et al., 2005). In the non-linear dimensionality-reduced UMAP space, cells from the samples stimulated with FGF4 (5 and 50 ng/ml) were well mixed, while cells without exogenous FGF4 were clearly separated (Figure 7B), indicating that the strongest changes in gene expression happened already between 0 and 5 ng/ml. Only very few cells grouped within the neighborhood of unstimulated cells, showing that most cells responded to FGF4 stimulation. This further excludes a change in the ratio between responsive and unresponsive cells as the primary mechanism to explain the graded response on the level of single genes when measured across cells (Figure 6).

Interestingly, the single-cell transcriptomes of the FGF-stimulated samples covered a larger space in the dimensionality reduction. This could indicate a higher variability in the single-cell transcriptomes. However, due to the non-linearity of UMAP, this cannot be concluded directly. To quantify if this correlates to a higher variability in mRNA counts per gene, the coefficient of variation of the expression of differentially expressed genes was compared between the samples. Higher FGF4 concentrations corresponded to higher median coefficients of variation (30.7% increase between 0 and 50 ng/ml FGF4, Figure 7C). The quantification of variation from single-cell RNA sequencing is subject to the caveat that noise levels per gene correlate with its expression levels at low expression levels. However, since only strongly expressed, up-, and downregulated genes were considered, higher median coefficients of variation are unlikely to reflect differences in technical noise levels between the samples. To estimate the contribution of extrinsic noise to the heterogeneity detected in the FGF-treated samples, the correlations between upregulated target genes among cells were calculated, and the coregulation of genes was depicted by hierarchical clustering (Figure 7D). In general, very low correlations were observed for most genes. Within one group of genes, including *Lef1* and *Tcf15*, the Pearson correlation coefficients reached 0.4, possibly indicating a set of cells with a gene expression change in Wnt-dependent genes. The absence of further, strongly coregulated groups of genes suggests the absence of functional heterogeneity and further specification into discrete cell states within the stimulated cells. The increase in variability, however, might be due to the stochasticity of events during gene transcription. In conclusion, stimulation with FGF4 increases the intrinsic noise and thus cells potentially explore greater regions of the high dimensional transcriptional space.

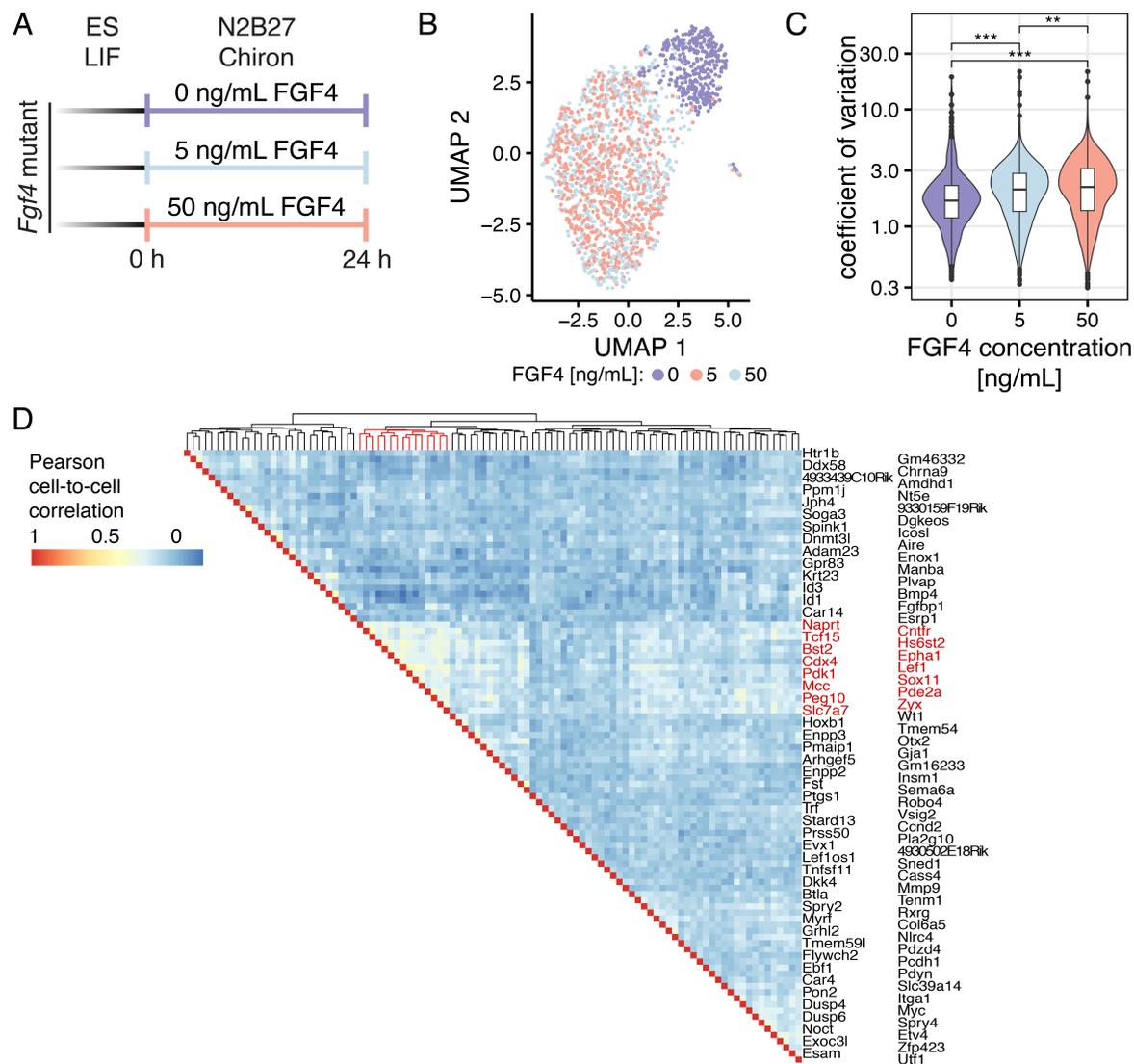


Figure 7: FGF-stimulation leads to increased gene expression variability in mESCs.

A Schematic of a single-cell RNA sequencing experiment to determine gene expression variability in mESCs upon FGF4 stimulation.

B UMAP representation of single-cell transcriptomes, under conditions from (A). Colors indicate FGF4 concentration.

C Heatmap of Pearson correlations between FGF-dependent genes (top 100 with highest expression fold change between 0 and 50 ng/mL FGF4 and p-value < 0.05) among cells treated with 50 ng/mL FGF4. ** indicates p < 0.01; *** indicates p < 0.001, Wilcoxon rank sum test.

D Coefficient of variation of all differentially expressed genes between 0 and 50 ng/mL FGF4 with a p-value < 0.05. Bonferroni adjusted Wilcoxon rank sum test. A possibly Wnt-dependent gene cluster is highlighted in red.

Single-cell RNA sequencing suggested that FGF influences the heterogeneity of gene expression. I next sought to characterize how FGF regulates the transcriptional bursting dynamics. I used the mRNA count distributions with single-cell resolution to estimate the transcriptional bursting kinetics and estimate the burst size and frequency (Figure 8A). Since the single-cell RNA sequencing data recapitulates only a fraction of the total mRNAs of a cell, for validation of the

obtained mRNA count distribution, and to estimate the capture and sequencing efficiency of the sequencing-based approach, I used in situ HCR to count mRNAs per cell for the previously identified FGF target genes *Spry4* and *Nanog* under the same experimental conditions as in the single-cell RNA sequencing experiment (Figure 7A, Figure 8B). For both genes, mRNA count distributions captured with both methods appeared similar in shape and reflected the supplemented FGF4 concentration (Figure 8C). The difference in sensitivity between in situ HCR and single-cell RNA-sequencing was gene-dependent. In the case of *Spry4*, one-ninth of the mRNAs were captured with single-cell RNA sequencing (Figure 5C, left), while for *Nanog*, the difference was 7.5-fold (Figure 8C, right). The differences between individual genes depend on both the in situ HCR efficiency as well as the capture efficiency during single-cell RNA sequencing. In general, the magnitude of the differences reflected well the combined factor of published capture efficiency of droplet-based single-cell RNA sequencing (30 – 32%; Zheng et al., 2017) and the proportion sequenced of the sequencing library at the reached sequencing saturation ($\approx 40\%$). Additionally, the factors aligned with the comparison of the detected mRNA counts for all genes per cell and the total mRNA content (Carter et al., 2005). The distributions of mRNA counts obtained from single-cell RNA sequencing methods were used to infer the burst size and frequency of target genes based on a profile-likelihood approach under the assumption of a steady state of gene transcription after 24 h of stimulation (Larsson et al., 2019). In the presence of FGF4, FGF target genes did not show a coherent difference in burst size compared to two randomly selected gene sets with similar expression strength. However, comparing the same gene sets, FGF target genes had a lower frequency of transcriptional bursts. This lower frequency is again reflected in a 14.5% higher median coefficient of variation compared to the control gene sets. In summary, FGF target genes, as genes that are strongly regulated during early differentiation, were expressed with higher variability between the cells, correlating to a generally lower transcriptional burst frequency within this gene set.

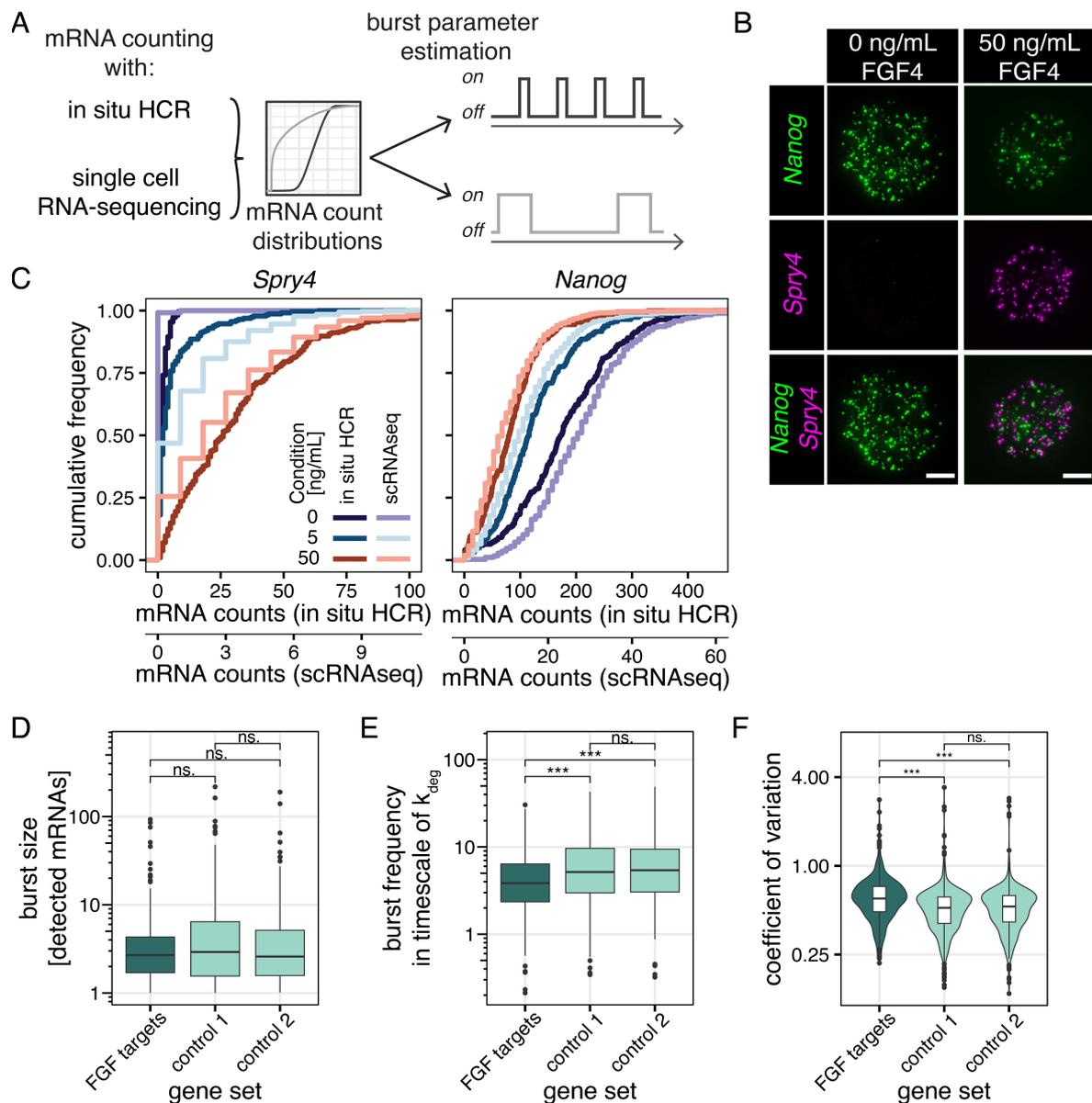


Figure 8: Reduced transcriptional burst frequency and increased cell-to-cell variation of transcription of FGF target genes.

A Schematic of mRNA counting for specific genes with in situ HCR and globally with single-cell RNA sequencing. Assuming a telegraph model of transcription, the transcriptional burst parameters burst size and burst frequency can be inferred from mRNA count distributions in dependency of the mRNA degradation rate.

B Images of *Nanog* and *Spry4* mRNAs within one cell under the indicated FGF signaling conditions.

C Cumulative mRNA count distributions for *Nanog* and *Spry4* from in situ HCR (dark colors) and single-cell RNA-sequencing (scRNAseq, bright colors) data. For in situ HCR, the distribution is representative for N = 2 replicates with n ≥ 200 cells per replicate.

D - F Estimated burst parameters burst frequency (D) and burst size (E) as well as the coefficient of variation (F) of all FGF target genes (log₂-fold change > |0.5|, adjusted p-value < 0.05 between 0 and 50 ng/mL FGF4 and minimal mean detection of 2.5 counts per cell) compared with two random sets of genes with the same profile of expression strength in the presence of 50 ng/mL FGF4. ns indicates p ≥ 0.05, *** indicates p < 0.001, Wilcoxon signed rank test.

4.2 Signaling and transcriptional regulators of *Spry4* expression

Transcriptomic analysis of the cellular response to FGF signaling showed substantial gene expression changes, defining up- and downregulated FGF target genes. These target genes were found to be more heterogeneously expressed between single cells in the same conditions, which correlated to lower transcriptional burst sizes. Next, I aimed to identify molecular mediators of such signaling induced transcriptional changes. For unbiased detection of these regulators, a genetic screen with a sensitive readout for the effects of FGF signaling was performed.

4.2.1 Establishing a *Spry4*^{H2B-Venus} reporter line

To track transcriptional activity in response to stimulation with FGF4, *Spry4* was shown to be a well-suited gene since it is strongly upregulated (Figure 6) and constantly expressed in FGF signaling conditions including Epi and PrE fates (Morgani et al., 2018). Previously, a reporter gene for *Spry4* was established, which did not hamper the developmental potential of the E14tg2a mESCs (Morgani et al., 2018). For compatibility with downstream experimental approaches, the reporter was re-established with its resistance changed from puromycin to neomycin (Figure 9A). These reporter cells showed nuclear H2B-Venus fluorescence upon release from 2i to ES + LIF, due to the paracrine FGF4 signaling (Figure 9B). Quantification by flow cytometry revealed a median 9.4-fold increase of H2B-Venus fluorescence in ES + LIF, resulting in a clear shift between the distributions in the two media conditions (Figure 9C).

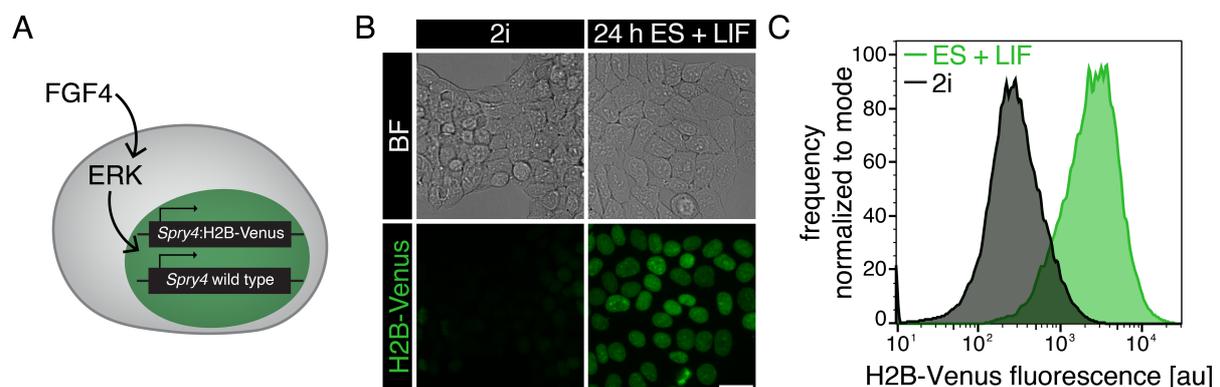


Figure 9: *Spry4*^{H2B-Venus/+} reporter cells as a readout for signaling-induced transcriptional activity.

A Schematic of the *Spry4*^{H2B-Venus/+} reporter cells.

B Images of *Spry4*^{H2B-Venus/+} reporter cells in inhibited (2i, left) and active paracrine FGF signaling (24 h in ES + LIF, right). Scale bar: 25 μ m.

C H2B-Venus expression analyzed with flow cytometry of *Spry4*^{H2B-Venus/+} reporter cells in 2i (black) and ES + LIF (green).

4.2.2 Genome-wide CRISPR knockout screen based on *Spry4*^{H2B-Venus} reporter activity

In order to identify signaling and transcriptional regulators required for signaling-induced mESC differentiation, I performed a genome-wide CRISPR knockout screen. The *Spry4*^{H2B-Venus/+} reporter cells and FAC-sorting were used as a sensitive readout for signaling-induced transcriptional activity. First, CAS9 was stably expressed in the reporter cells to ensure that subsequent transduction with lentiviral particles containing a pooled gRNA library results in efficient knockout of targeted protein-coding genes (Figure 10). To test the influence of all protein-coding genes on *Spry4*^{H2B-Venus/+} reporter activity, I used the pooled gRNA library Brie, which includes 78,637 gRNA sequences targeting 19,674 different genes. With flow cytometry, reporter cells with the highest and lowest *Spry4*^{H2B-Venus} fluorescence in ES + LIF medium were sorted. To identify positive and negative regulators, the 1% and 5% lowermost and topmost cells were analyzed at two different time points after gRNA library transduction to estimate robustness over passaging of corresponding knockout cells. From all sorted populations, enriched gRNAs were determined by amplification and sequencing of genome-integrated gRNAs and by comparing them to a non-sorted control sample from the respective time point (Figure 10).

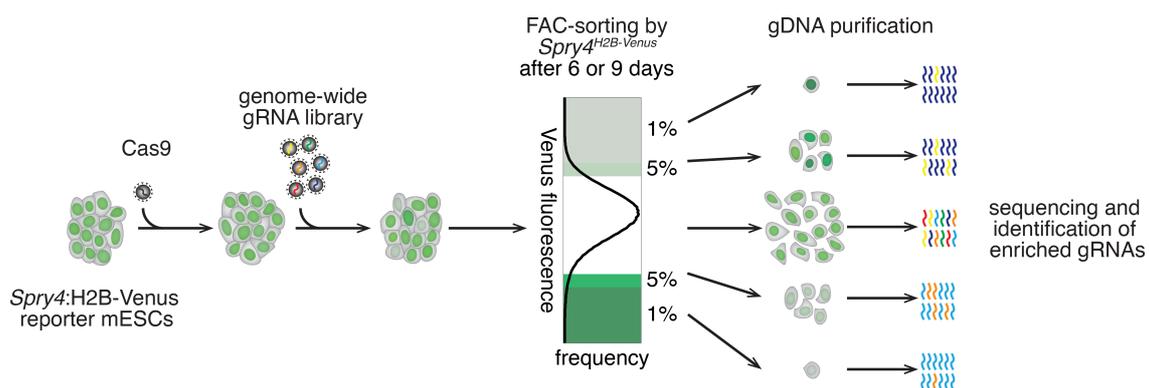


Figure 10: Schematic of a genome-wide CRISPR knockout screen.

Cas9-expressing *Spry4*^{H2B-Venus/+} reporter cells were transduced with a pooled gRNA library targeting protein-coding genes (Doench et al., 2016). FAC-sorting for cells with in- or decreased H2B-Venus fluorescence was performed after 6 and 9 days. gRNAs enriched in sorted fractions were identified by sequencing.

To assess whether the complexity of the library was preserved during transduction, cell expansion and splitting (see chapter 3.4), the distribution of gRNA counts was plotted for each sample and control condition (Figure 11A). The library control showed a unimodal distribution with only a minimal number of non-detected gRNAs. In both non-sorted conditions, a few more gRNAs, most likely targeting essential genes, are underrepresented, with no substantial differences between days 6 and 9 after transfection. Among the gRNAs that lost representation most strongly, many target different ribosomal subunits (Supplementary Figure 1), and thereby,

the knockout of those genes led to limited proliferation rates. In contrast, for example, *Trp53*, a tumor suppressor gene known to regulate the cell cycle at the G1/S regulation point (ter Huurne et al., 2020), is overrepresented in transduced cells compared to the library control (Supplementary Figure 2).

In the sorted populations, a high number of gRNAs were only detected at background levels (gRNA counts < 100). The number of detected gRNA with higher counts decreased with time after transduction and when comparing 1% versus 5% gates (Figure 11A). This reflected the specificity of the sorting approach enriching gRNAs truly affecting *Spry4^{H2B-Venus}* expression. In the genome-wide library, 1000 non-targeting control gRNAs were included (Doench et al., 2016) to further estimate the variability in the enrichment in sorted population. In all sorted conditions, a number of gene-targeting gRNAs had higher fold changes than non-targeting gRNAs (Figure 11B). Combining gRNAs targeting the same gene and ranking genes with the robust-rank-algorithm (W. Li et al., 2014) revealed up to 17 individual genes with an FDR ≤ 0.05 and up to 26 genes with an FDR < 0.2 enriched in cell populations with the lowest *Spry4^{H2B-Venus}* expression (Figure 11C, left). In cells with the highest *Spry4^{H2B-Venus}* expression, up to 12 and 24 genes with FDR values of < 0.05 and < 0.2, respectively, were found (Figure 11C, right). Together, the specific enrichment of targeting over non-targeting gRNAs, the identity of enriched and depleted genes, and the limited number of significantly enriched genes showed the specificity of the screening approach to robustly identify regulators of *Spry4^{H2B-Venus}* reporter expression.

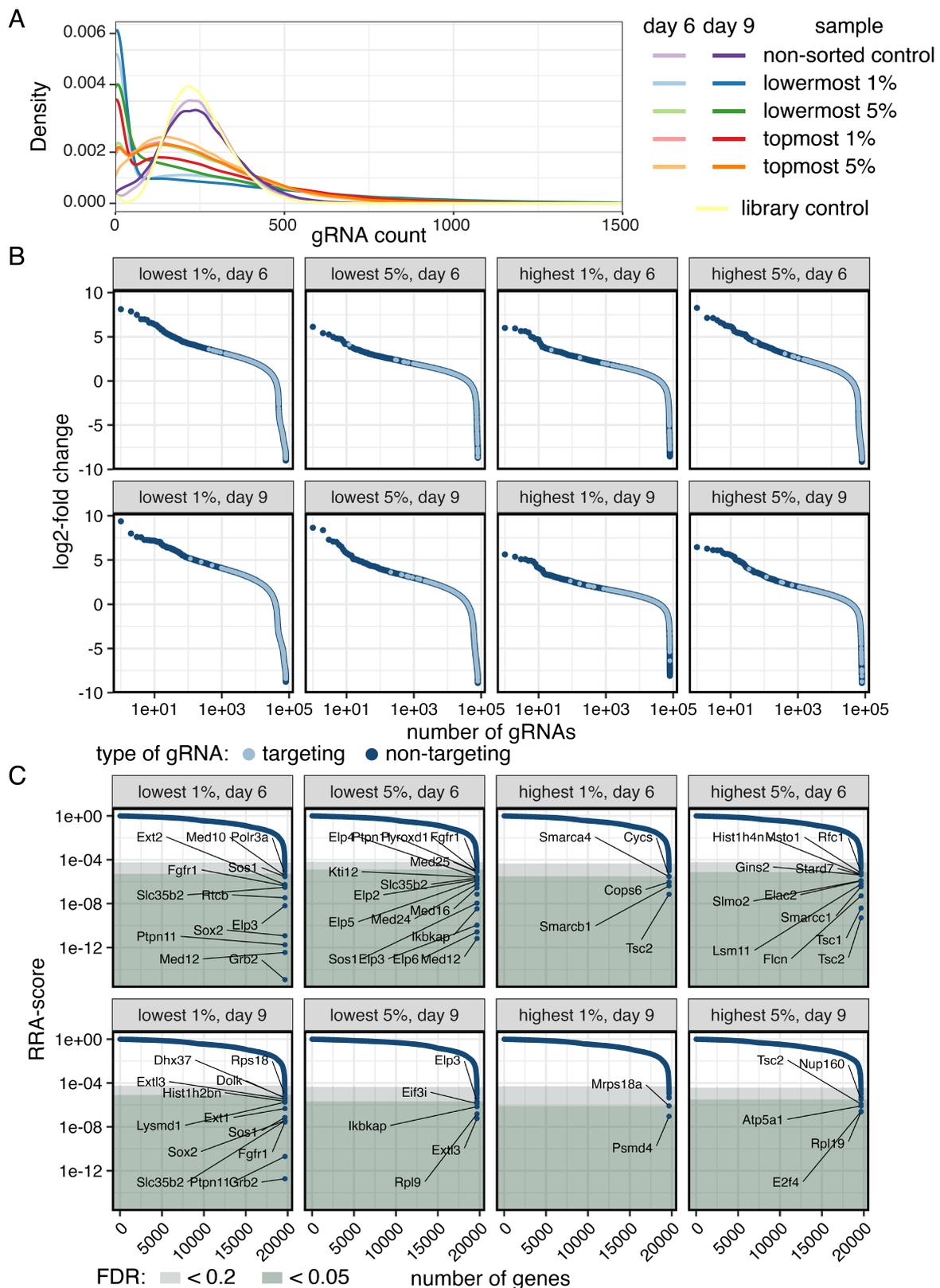


Figure 11: Enrichment of gRNAs and corresponding genes.

A gRNA count distributions in the indicated samples and controls detected by sequencing.

B,C Enrichment of gRNAs (B) and corresponding genes (C) in cells sorted for the lowermost 1% of H2B-Venus signal 6 days after gRNA transduction (left), the lowermost 5% of H2B-Venus signal 6 days after gRNA transduction (second from left), the topmost 1% of H2B-Venus signal 6 days after gRNA transduction (third
(Continued on next page)

from left) and the topmost 5% of H2B-Venus signal 6 days after gRNA transduction (first from right). Displayed as log₂-fold change (B) or RRA score of corresponding genes (C). In (B), gene-targeting gRNAs are in dark blue, and control gRNAs are in light blue. Green and gray background in (C) indicate FDR < 0.05 and < 0.2, respectively. All genes with an FDR < 0.05 are labelled.

4.2.3 Positive regulators of *Spry4* expression

First, I focused on gene perturbations identified in the screen leading to decreased *Spry4*^{H2B-Venus} expression. Since many genes reoccurred in more than one of the sorted conditions (Figure 11C), I limited my further analysis to genes that either have an FDR < 0.05 in one of the conditions or an FDR < 0.2 in at least two conditions. Hierarchical clustering based on the FDR values from all sorting conditions showed groups of the 35 genes with different effect sizes and reproducibility between the sorting time points and gates (Figure 12A). Genes with presumably very strong effects were only enriched in the more stringent 1% gates (e. g. *Grb2*), whereas subtle effects resulted in enrichment in the 5% gates (e. g. *Mgat1*). Gene knockouts only significantly enriched in the later timepoint might indicate an extended period for the knockout to become functional or the underrepresentation of the corresponding gRNAs in the non-sorted population due to a growth disadvantage, for example, for *Histh2bn* or *Dhx37* (Figure 12A, Supplementary Figure 3). A protein interaction graph was constructed to corroborate further the relation between the resulting genes (Figure 12B) based on known interactions and associations from StringDB (Szklarczyk et al., 2015). This database includes experimentally proven protein-protein interactions and bioinformatically predicted interactions, as well as assumed functional relations based on, for example, the mentioning of two protein names together in research paper abstracts. Unifying this information, most hits fell in larger, highly connected groups of genes. These groups could be manually annotated with molecular functions based on a known cellular function of the individual genes (Figure 12A). Non-connected hits were partially included in these groups when known functions fitted that of preexisting groups. With this approach, five main clusters were identified: First, the finding of genes directly positively influencing the FGF signaling cascade *Fgfr1*, *Grb2*, *Sos1*, and *Ptpn11* confirmed the strong dependence of *Spry4* on active FGF signaling. Inefficient FGF signaling most likely also explains the enrichment of a group of genes (*Slc35b2*, *Ext1*, *Ext2*, and *Extl3*) involved in protein glycosylation and, specifically, the synthesis of heparan sulfates. Heparane sulfates are important co-factors for FGF binding and signal perception at FGF receptors (Ornitz & Itoh, 2015). The occurrence of these genes in the pooled, genome-wide design of the CRISPR knockout screen, in which the number of neighbor cells with the same genotype is limited by cell division and cell reseeded, points to a cell-autonomous function of heparan sulfates. Heparane sulfate synthesis thereby constitutes an effective mechanism to regulate a cell's responsiveness to external FGF signals. Two groups of enriched genes included subunits of two

complexes possibly related to transcription: Elongator and Mediator. For Elongator, six subunits (*Ikkkap*, *Elp2* to *Elp6*) and its closely related protein KTI12 were identified to positively affect *Spry4* expression. It remains unclear if Elongator regulates the transcription of the *Spry4*^{H2B-Venus} locus or if the regulation is further downstream during translation (Hermand, 2020; Svejstrup, 2007). Mediator hits included five subunits from the middle (*Med10*), tail (*Med16*, *Med24*, *Med25*), and kinase modules (*Med12*, Yin & Wang, 2014). The last group of genes included hits among core ribosomal and translational functions (*Rpl9*, *Rpl18*, *Eif3i*, and *Dhx37*). RNA splicing-related hits included *Fam98b* and *Rtcb*. The loss of the pluripotency-associated transcription factor *Sox2* might have influenced reporter gene expression by inducing further differentiation, since *Sox2* is an important pluripotency factor. *Hist1h2bn*, rather than being regulated explicitly at the endogenous *Spry4* locus, is more likely to only affect the *Spry4*^{H2B-Venus}-reporter locus since only gRNAs targeting sequences present in the H2B-targeting construct were significantly enriched. Thus, the CRISPR knockout screen confirmed known signaling regulators (components of the FGF signaling system and heparin sulfate synthesis) and identified new potential transcriptional and translational regulators (e. g. Mediator and Elongator subunits) of *Spry4* expression in ESCs (Figure 12A,B).

To further test the reproducibility of the found hits and their magnitude of effect, I used polyclonal knockout lines in the *Spry4*^{H2B-Venus/+}-reporter background. Targeted genes for validation (Figure 12B, indicated by a black circle) were selected by their effect sizes in the screen, their representation in the non-sorted conditions to eliminate cell toxicity (Supplementary Figure 3) and limited to a few subunits of the same protein complex or affected signaling system. This list included subunits of the Elongator and Mediator complexes, *Grb2* and *Ptpn11* as FGF signaling factors and *Sox2* and *Fam98b*, which were not clearly linked to a functional group. I analyzed *Spry4*^{H2B-Venus} expression with flow cytometry under similar conditions as in the screen 6 days after transfection in ES + LIF. *Spry4*^{H2B-Venus} expression was significantly reduced in all tested mutants (Figure 12C,D). Mutation of *Grb2* and *Ptpn11* lead to strong reduction in *Spry4*^{H2B-Venus} expression. However, comparison with the effect of MAPK inhibitor treatment with PD03 demonstrated an even slightly stronger effect of the complete loss in FGF/ERK activity, indicating either redundancy mechanisms or incomplete knockout in the polyclonal cells. All other perturbation led to *Spry4*^{H2B-Venus} expression levels between 35% and 85% of the parental wild-type line. While most Mediator subunits showed a mild reduction in *Spry4*:H2B-Venus levels, the Mediator subunit *Med12* stood out with having the strongest effect, reducing *Spry4*:H2B-Venus levels to 35.3% ± 1.6% compared to control. Together, the mutation of newly identified transcriptional regulators robustly lowers *Spry4*^{H2B-Venus} reporter levels but does not entirely reflect the loss of all FGF signaling activity. The consistent reduction in *Spry4*^{H2B-Venus} reporter levels for all tested hits emphasizes the sensitivity of the screening approach and points out *Med12* as the transcriptional

regulator most strongly influencing reporter levels. Given the hits among the FGF signaling cascade, the screen suggests *Med12* as a candidate transcriptional regulator itself regulated by signaling.

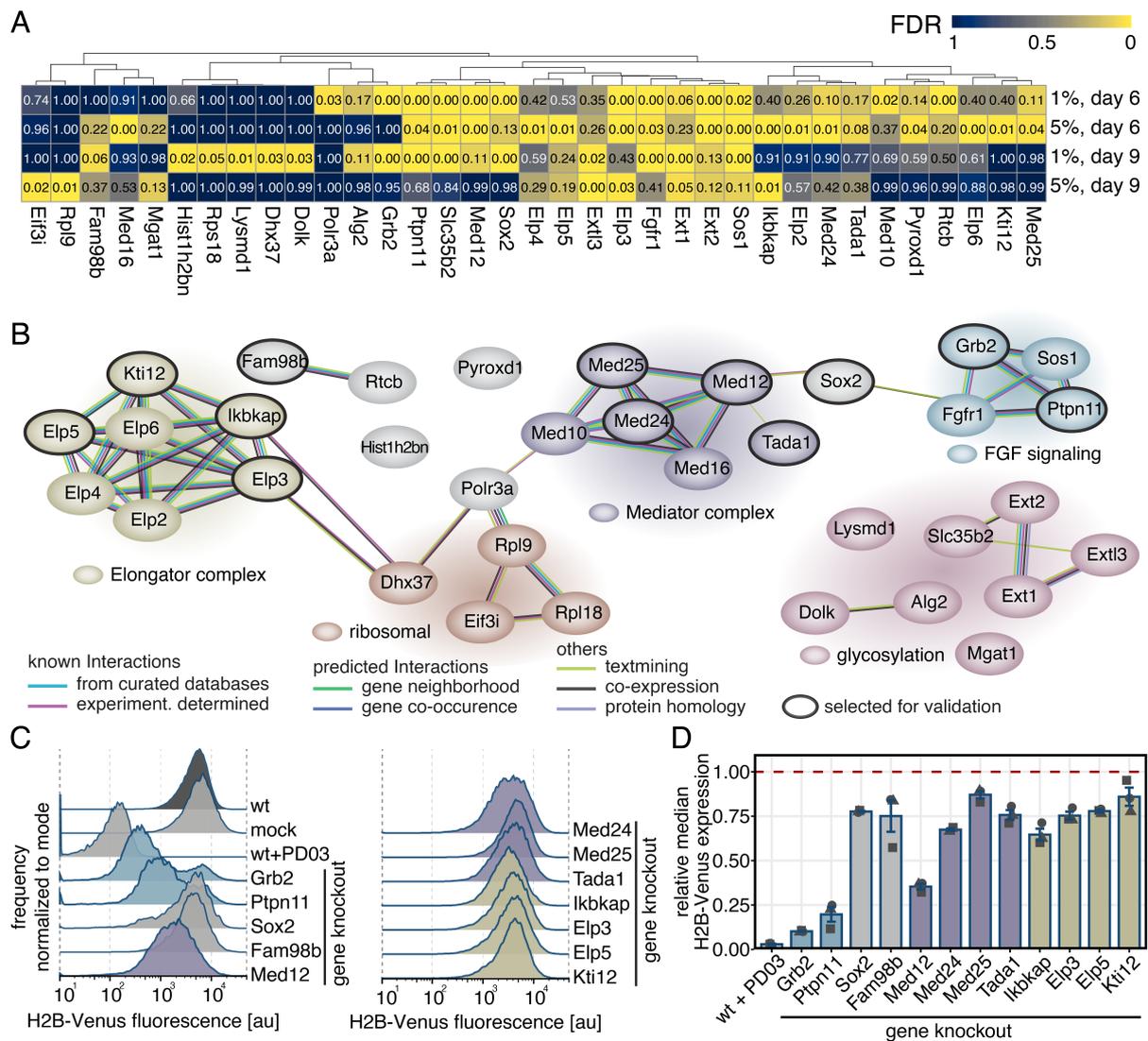


Figure 12: Genome-wide CRISPR knockout screen reveals positive regulators of *Spry4* expression.

A Hierarchical clustering of gene perturbations leading to reduced *Spry4*^{H2B-Venus} expression (FDR < 0.05 in at least one condition or FDR < 0.2 in at least two conditions).

B Protein Interaction Network of genes shown in (A) is based on String-DB. Background colors of genes and gene clusters were manually assigned based on classification by functional similarity.

C,D H2B-Venus expression in *Spry4*^{H2B-Venus/+} reporter cells upon knockout of selected candidate genes 6 d after transfection. (C) shows histograms of one representative experiment, (D) shows mean \pm SEM of median *Spry4*^{H2B-Venus} expression from N = 3 independent experiments. $p < 0.05$ for wild type vs. *Elp3*, *Elp5*, *Fam98b*, *Ikbkap*, *Kti12*, *Med25* or *Tada1* knockouts; $p < 0.01$ for wild type vs. *Grb2*, *Med12*, *Med24*, *Ptpn11*, *Sox2* knockouts or PD03-treated cells (Benjamini-Hochberg-adjusted, one-sided, paired t-test).

4.2.4 Negative regulators of *Spry4* expression

Although *Spry4*^{H2B-Venus} is expressed at constant high levels in ES + LIF medium, negative feedback mechanisms could prevent even stronger expression. To find negative regulators of transcription of *Spry4*, I analyzed the enriched gRNA-mediated knockouts in the cells with the highest *Spry4* expression. Again, genes with an FDR < 0.05 or genes with an FDR < 0.2 in at least two sorting conditions were selected for hierarchical clustering based on the FDR values (Figure 13A). 29 genes were identified and grouped into functional classes with the help of their interaction partners (Figure 13B). Again, an FGF signaling-related gene, *Lztr1*, was among the enriched genes. LZTR1 functions as an adapter to ubiquitinylate and thereby inhibit RAS and downstream mitogen-activated protein kinase activity (Steklov et al., 2018), in line with the overrepresentation of positive regulators of FGF signaling in the cells showing the lowest *Spry4*^{H2B-Venus} levels. The previously described growth advantage in other cell types (Steklov et al., 2018) was not confirmed by any increased representation in the non-sorted conditions (Supplementary Figure 5). With *Rpl19* and *Elac2*, two genes are related to translation in general, contrasting the ribosomal subunit knockouts in the cells with low *Spry4*^{H2B-Venus} expression. A large group of genes related to the electron transport chain and other mitochondrial functions suggest a high sensitivity of *Spry4*^{H2B-Venus} expression on the cell's metabolic state. The enrichment of the chromatin remodelers *Smarcc1*, *Smarcb1*, and *Smarca4*, as well as the chromatin-associated proteins *Hist1h4n*, *Trrap*, and *E2f4*, suggest that the chromatin around the *Spry4* locus is actively regulated for its expression. Four genes associated with mTOR signaling (*Tsc1*, *Tsc2*, *Fln*, and *Lamtor4*) were identified as the last group associated with cell signaling and the metabolic state of the cell. Together, similar to the previously identified positive hits, the screen revealed regulators that exert negative control over *Spry4* expression. These include proteins influencing *Spry4* expression by negatively influencing signaling potentially upstream of *Spry4* and proteins inhibiting the transcription of *Spry4* expression more directly by chromatin modifications.

Again, I aimed to validate and measure the effect sizes of the hits by analyzing *Spry4* expression in polyclonal knockout lines for a selected range of identified genes focusing on mTOR signaling, FGF signaling, and chromatin regulation. Deletion of each of the six candidate genes using single gRNAs resulted in a significant increase of mean *Spry4*^{H2B-Venus} fluorescence levels in reporter cells. The impact of knocking out the mTOR signaling genes surpassed that of targeting *Lztr1* or *Smarcc1*, more than doubling reporter expression levels compared to mock-transfected cells (Figure 13C,D). To summarize, components regulating mTOR activity had the most substantial negative effect on *Spry4*^{H2B-Venus} expression and suggest a role of mTOR signaling in cell state-specific gene expression.

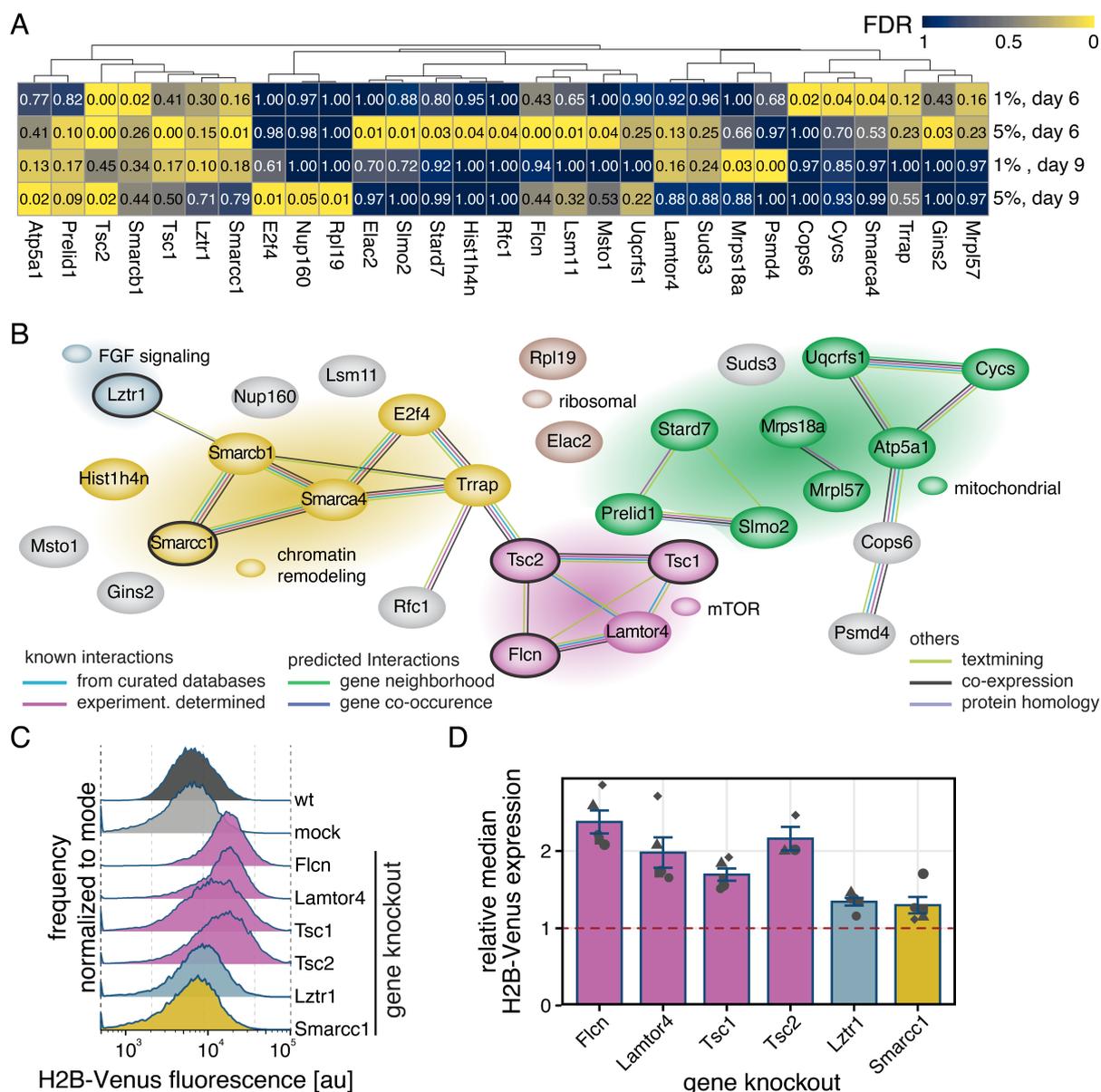


Figure 13: Genome-wide CRISPR knockout screen reveals negative regulators of *Sprouty4* expression.

A Hierarchical clustering of gene perturbations leading to increased *Spry4*^{H2B-Venus} expression (FDR < 0.05 in at least one condition or FDR < 0.2 in at least two conditions).

B Protein Interaction Network of genes shown in (A) based on String-DB. Background colors of genes and gene clusters were manually assigned based on classification by functional similarity.

C,D H2B-Venus expression in *Spry4*^{H2B-Venus/+} reporter cells upon knockout of selected candidate genes 6 d after transfection. (C) shows histograms of one representative experiment, (D) shows mean \pm SEM of median expression from $n = 3$ independent experiments. $p < 0.05$ for mock-transfected wild type vs. *Smarcc1* knockout; $p < 0.01$ for mock-transfected wild type versus all other knockouts (Benjamini-Hochberg-adjusted, one-sided, paired t-test).

4.3 *Med12* modulates gene expression independently from signaling in pluripotency

4.3.1 Establishment of *Med12* mutant lines

Next, I sought to find out how the transcriptional regulators identified in the screen relate to pluripotency-associated signaling systems. Multiple Mediator subunits appeared to be positive regulators of *Spry4*. The polyclonal knockout of *Med12* had the strongest effect on the transcriptional reporter expression of all Mediator subunits, indicating a critical role of this subunit in order to react to extrinsic signals. Moreover, other identified subunits showed a loss over time in the transduced cells (Supplementary Figure 5), indicating a more general and essential role of the different modules of the Mediator complex. The kinase module (Lynch et al., 2020) and *Med12* itself (Rocha et al., 2010) have previously been linked to pluripotency-associated signaling systems. Therefore, I investigated the relation between multiple of these signaling systems and the function of *Med12* in mESC differentiation.

On that account, I created cell lines with a deletion in the seventh exon of the *Med12* gene (Figure 14A), localized on the X-chromosome. A combination of two gRNAs, targeting the same region that resulted in the strongest effect in the CRISPR knockout screen, was used in three different parental lines: The *Spry4^{H2B-Venus}* reporter line, an *Fgf4*-mutant line (see chapter 3.2) to control FGF signaling precisely, and a *Gata6*-inducible line for further experiments analyzing PrE differentiation (see chapter 3.2 and 4.5). Western blot analysis confirmed the loss of MED12 in the clonal lines (Figure 14B). In the *Spry4^{H2B-Venus/+}*-reporter line, the loss of *Med12* resulted in a decreased reporter gene expression upon release from 2i + LIF into N2B27 (Figure 14C,D), confirming the results from the screen and polyclonal knockout lines (Figure 12). The *Med12*-mutant lines proliferated normally and looked morphologically similar to wild-type cells (Figure 14D). To confirm that *Med12* influences *Spry4^{H2B-Venus}* levels by regulating transcription, phospho-ERK levels between wild-type and *Med12*-mutant cells were compared. In different backgrounds, there was no consistent difference in phospho-ERK levels relative to total ERK levels (Figure 14E,F), showing that the decreased *Spry4^{H2B-Venus}* reporter levels were not due to decreased phospho-ERK levels and suggesting a role of *Med12* as a transcriptional regulator. These *Med12* loss-of-function mESC lines will be referred to as *Med12* mutants and serve as an important tool in further experiments to resolve the role of *Med12* in pluripotency and differentiation.

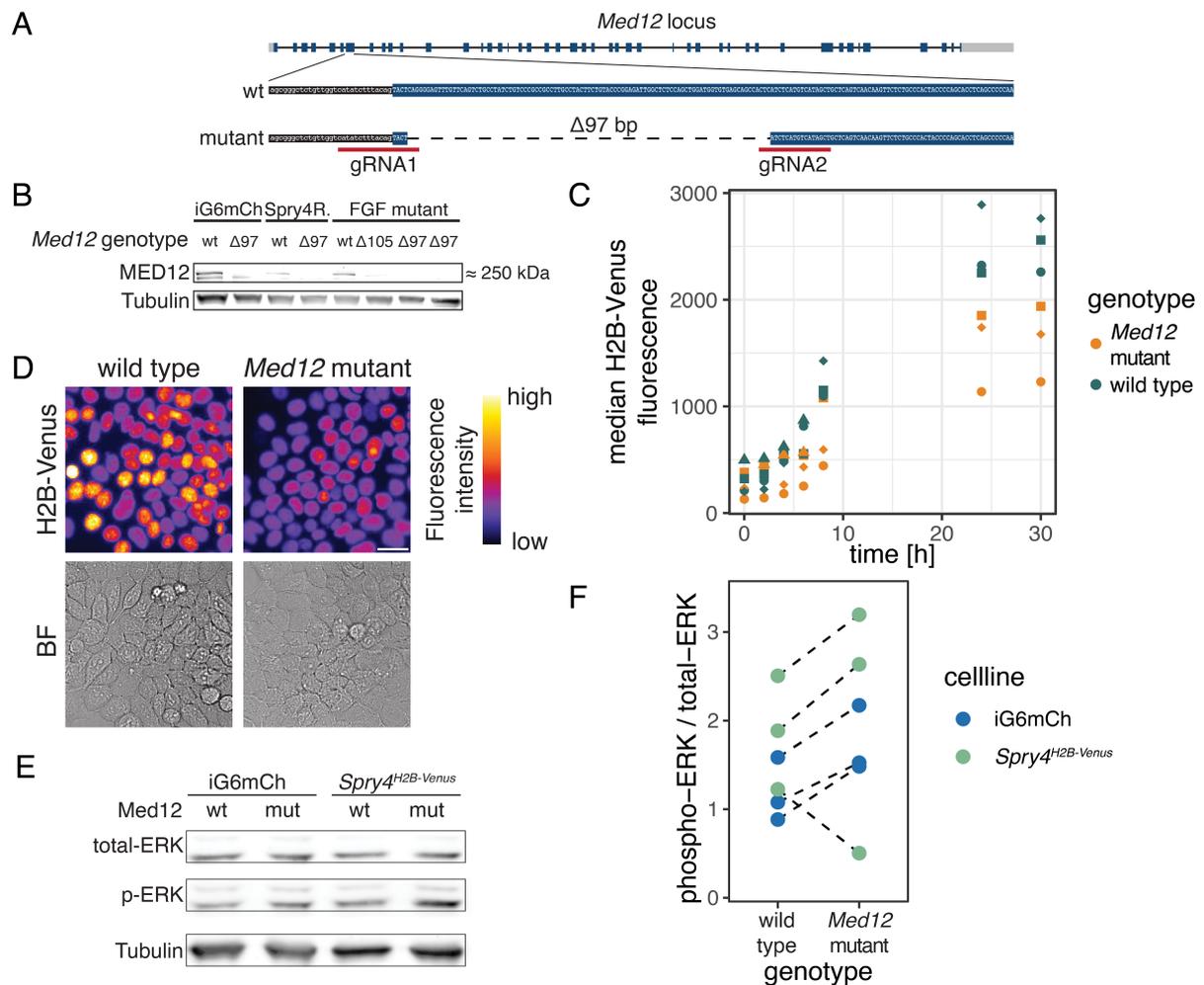


Figure 14: Generation of *Med12* mutant cell lines.

A Schematic of the *Med12* gene locus and the gRNAs used to create a *Med12* loss-of-function by deleting part of exon 7.

B Immunoblotting of cell lysates from several monoclonal *Med12* mutant lines generated in different genetic backgrounds, stained for MED12 and Tubulin.

C *Spry4^{H2B-Venus/+}* expression upon release from 2i + LIF to N2B27 in wild-type and *Med12*-mutant cells measured by flow cytometry. Data points show median fluorescence in each experiment indicated by point shape. N = 3.

D H2B-Venus expression in live wild-type and *Med12*-mutant *Spry4^{H2B-Venus/+}* cells after 24 h of growth in N2B27 following release from 2i + LIF.

E Immunoblotting of cell lysates from *Med12* wild type and *Med12*-mutant *Spry4^{H2B-Venus/+}*-Reporter and iGata6 mESCs, stained for Tubulin, total- and phospho-ERK.

F Quantification of phospho-ERK signals from immunoblots, normalized to total-ERK. N=3.

4.3.2 Loss of *Med12* affects gene expression in pluripotency and differentiation

Med12 was highlighted as a positive regulator of *Spry4* expression in the CRISPR screen and validation of poly- and monoclonal knockout lines. I next asked to how many and to which genes besides *Spry4* the function of *Med12* is extended to during pluripotency and differentiation of mESCs. Utilizing the *Med12*-mutant line, gene expression was compared to the wild type with a

bulk RNA sequencing experiment. mRNA abundances were obtained for both genotypes in pluripotent cells in 2i as well as in cells differentiated in N2B27 for 24 h (Figure 15A). Principal component analysis showed that the replicates were separated by their experimental conditions and genotypes (Figure 15B). The first principal component, explaining 32.1% of the total variability in the dataset, distinguishes pluripotent from differentiating conditions. The second principal component corresponded to 16.8% of the variability and separated the samples by their genotype, indicating that cells transitioning from pluripotency to differentiation are transcriptionally more distinct from each other than wild-type and *Med12*-mutant cells (Figure 15B). This was further confirmed by the number of differentially expressed genes comparing the two genotypes or media conditions. A higher number of genes were significantly differentially expressed upon differentiation than between the genotypes (Figure 15C). Moreover, 88% of the genes having a log₂-fold change greater than |1| during differentiation in *Med12*-mutant cells met the same criterium in wild-type cells (Figure 15C, left), indicating that most *Med12*-mutant cells follow the same differentiation trajectory as the wild-type cells. On the other hand, the number of genes differentially expressed between wild-type and *Med12*-mutant cells increases from 307 to 423 genes during the transition from pluripotency in 2i to differentiation in N2B27. 163 of these genes (40.3%) are differentially expressed in both media conditions (Figure 15C, right), suggesting that the transcriptomic states of wild-type and *Med12*-mutant cells are more similar in pluripotency than upon differentiation. As a first approach to finding out how wild-type and *Med12*-mutant cells are different in terms of their differentiation behavior, I analyzed the enriched biological functions associated with the differentially expressed genes using the gene ontology resource (Aleksander et al., 2023; Ashburner et al., 2000). In pluripotency, the list of enriched biological functions was dominated by different levels of regulation of the cytoskeleton, particularly related to actin (Figure 15D). Further higher-level biological functions enriched significantly in the known functions of the differentially expressed genes between the genotypes included morphogenesis and cell motility. Many of these are still significantly enriched after differentiating wild-type and *Med12*-mutant cells. Additionally, in differentiation conditions, multiple functions related to differentiation have a low FDR (Figure 15E). Notably, these included terms related to neural differentiation (neuron development, neuron projection development, generation of neurons, neurogenesis, neuron differentiation) and placenta development (embryonic placenta development and placenta development). The differential gene expression in genes that have annotated function in neural development was in line with a potential differentiation phenotype of the *Med12*-mutant cells since wild-type cells differentiate towards a neural cell identity upon culturing in N2B27 for more extended periods of time (Ying et al., 2003).

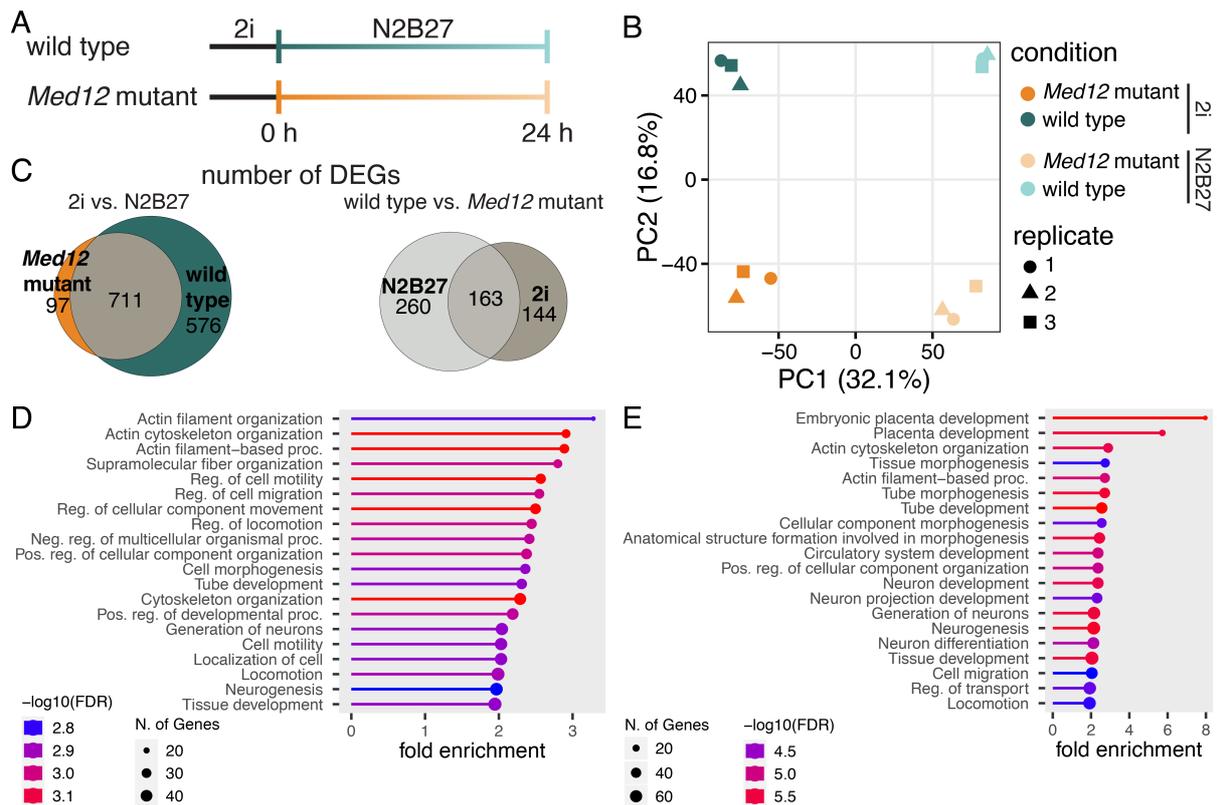


Figure 15: Differentially expressed genes between wild-type and *Med12*-mutant cells in pluripotency and differentiation.

A Schematic of experiment to identify *Med12*-regulated genes by bulk RNA sequencing.

B Principal component analysis transcriptomes from (A).

C Euler-diagram showing the number of differentially expressed genes (\log_2 -fold change $> |1|$, adjusted p-value < 0.01) in bulk transcriptomes. Left panel compares genes differentially expressed upon 24 h of differentiation between *Med12*-mutant and wild-type cells, right panel compares genes differentially expressed upon loss of *Med12* between N2B27 and 2i.

D,E Enriched GO-terms (biological functions) associated with differentially expressed genes between wild-type and *Med12*-mutant cells in 2i (D) and N2B27 (E). Differentially expressed genes were defined as in (C).

4.3.3 *Med12* regulates gene expression independently from pluripotency-associated signals

Mutation of *Med12* led to subtle changes in gene expression in pluripotency, which became more prominent during differentiation. It is unclear if this is due to an inability to implement a specific differentiation signal or if the transcriptional response of *Med12*-mutant cells is generally affected. To find out if any early differentiation-associated signaling system is affected, I compared the effects of the *Med12* loss of function to other signaling gene knockouts. I used a previously published dataset containing transcriptomes of more than 70 different mutants during differentiation (Lackner et al., 2021). Within these 70 mutants, central regulators of exit of pluripotency-associated signaling systems (Betschinger et al., 2013; Kalkan et al., 2019; Ying et al., 2008) were used to define a set of expression footprint genes for mTOR, Notch, Wnt and FGF

signaling. Their expression changes during differentiation were compared between different knockout and wild-type cells (Lackner et al., 2021). I included the *Med12*-mutant line in this analysis (Figure 16). The expression footprint during differentiation of the *Med12*-mutant line was most similar to knockouts of mTOR signaling components. Limited similarity was observed for Wnt and Notch footprint genes, while the expression changes in FGF-signaling mutants did not correlate with the expression changes in the *Med12*-mutant cells.

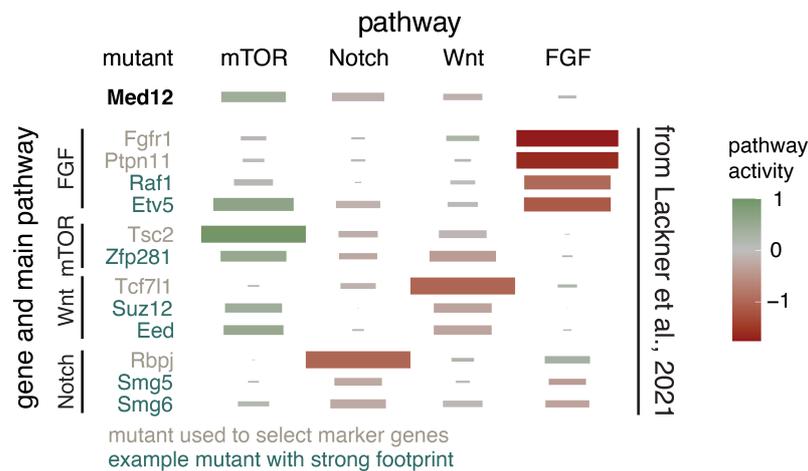


Figure 16: *Med12* mutant cell's expression footprint most similar to mTOR perturbation.

Expression footprint analysis using a set of 50 marker genes per pathway defined in (Lackner et al., 2021). The top row shows the footprint of *Med12*-mutant cells, and the lower rows show the footprint of pathway-defining mutants (gray) and example mutants (green) from Lackner et al., 2021, for comparison. Tile color indicates relative pathway activity, and the tile size indicates the Spearman correlation of footprint genes with pathway-defining mutants.

From this expression footprinting analysis mTOR signaling-mediated gene regulation appeared to be affected by the loss of MED12. Additionally, opposing to *Med12*, mTOR regulating genes inhibited *Spry4^{H2B-Venus}* expression in the CRISPR screen, suggesting an inhibiting causality between *Med12* and the mTOR regulating genes. To investigate the relationship between mTOR signaling and *Med12*-regulated transcription, I transfected *Spry4^{H2B-Venus}* reporter cells with the gRNAs targeting the mTOR regulators revealed by the screen *Tsc1*, *Tsc2*, *Flcn*, and *Lamtor4*. Similar to the validation experiment described in Figure 13D, 6 days after transfection in ES + LIF analysis of the reporter expression showed that *Spry4^{H2B-Venus}*-levels are still upregulated upon mTOR regulator knockout, even in the absence of MED12, though from a lower baseline level (Figure 17A). Even the relative effect sizes between the knockouts (Figure 17A, background) were conserved in the *Med12*-mutant background, contradicting the footprint analysis and indicating that *Med12* acts independently of mTOR signaling on the expression of *Spry4^{H2B-Venus}*.

To test if the proposed dependency of *Med12* on Wnt signaling activity in early development (Rocha et al., 2010) caused the reduced *Spry4^{H2B-Venus}* expression in the CRISPR screen and the

validation experiments, I analyzed the effect of Wnt inhibition and activation via small molecule treatment on *Spry4*^{H2B-Venus} expression in both wild-type and *Med12*-mutant lines. Neither the Wnt agonist Chiron nor the Wnt inhibitor XAV939 (XAV) influences *Spry4*^{H2B-Venus} expression in wild-type or *Med12*-mutant cells (Figure 17B). This argues for a Wnt-independent role of *Med12* in transcriptional regulation during pluripotency and transition to differentiation.

Another signaling system important during early differentiation is FGF-signaling. Recently, the mediator kinase module, including its subunit MED12, has been proposed to function downstream of FGF-mediated MAPK signaling (Lynch et al., 2020), and the CRISPR screen revealed hits within the signaling cascade. In contrast, the footprinting analysis did not suggest a strong correlation between how *Med12* and FGF affect gene regulation. To further resolve these results, I compared transcriptomes of *Med12* wild-type with mutant cells in an FGF-mutant background, allowing to control FGF signaling by stimulation with exogenous FGF4 (Figure 17C). Bulk RNA-sequencing was performed 6 h after transfer to N2B27 with or without FGF4 to capture immediate gene regulation. Again, principal component analysis grouped replicates in proximity in the first principal components, including multiple clonal lines of the *Med12* mutant (Figure 17D). Samples were segregated by their FGF stimulation state, explaining 33.2% of the total variance between the samples and by their genotype (PC3, 9.3% of variance). If the expression of FGF target genes predominantly relied on *Med12*, their gene expression fold change upon FGF stimulation would be expected to be lower in *Med12*-mutant cells than in wild-type cells. However, when visualizing the gene expression fold change for each gene upon 6 h of FGF stimulation in wild-type versus *Med12*-mutant cells, the majority of genes were induced to a similar degree in both genotypes (2006 genes), while only 926 and 935 genes were differentially expressed in the wild-type or the *Med12*-mutant only, respectively (Figure 17E). The few genes showing a clear dependency on *Med12* (*Lgals7*, *Gm32200*, *Tmem171*, *Cyth4*, *Gm30692*, *Tmprss2*, *Gm10371*, *Gm10648*, *4930598F16Rik*, *Ajap1*) had diverse or unknown function and showed no prominent common characteristics. Upregulated FGF target genes are still upregulated in the *Med12*-mutant cells with a median fold change ratio between wild-type and *Med12*-mutant cells of 0.927. Similarly, this median ratio was 1.059 for all downregulated genes. In both directions, the distributions of the ratio were unimodal with long tails, indicating high variability from gene to gene but no separate classes of FGF target genes in terms of their *Med12* dependence. By introducing a minor shift in the median away from 1, it remains a possibility that Med12 mildly influences the magnitude of FGF target genes. Nevertheless, these findings do not support a robust and distinct involvement of *Med12* in the control of FGF target genes, as well as in mTOR- or Wnt-dependent transcriptional regulation.

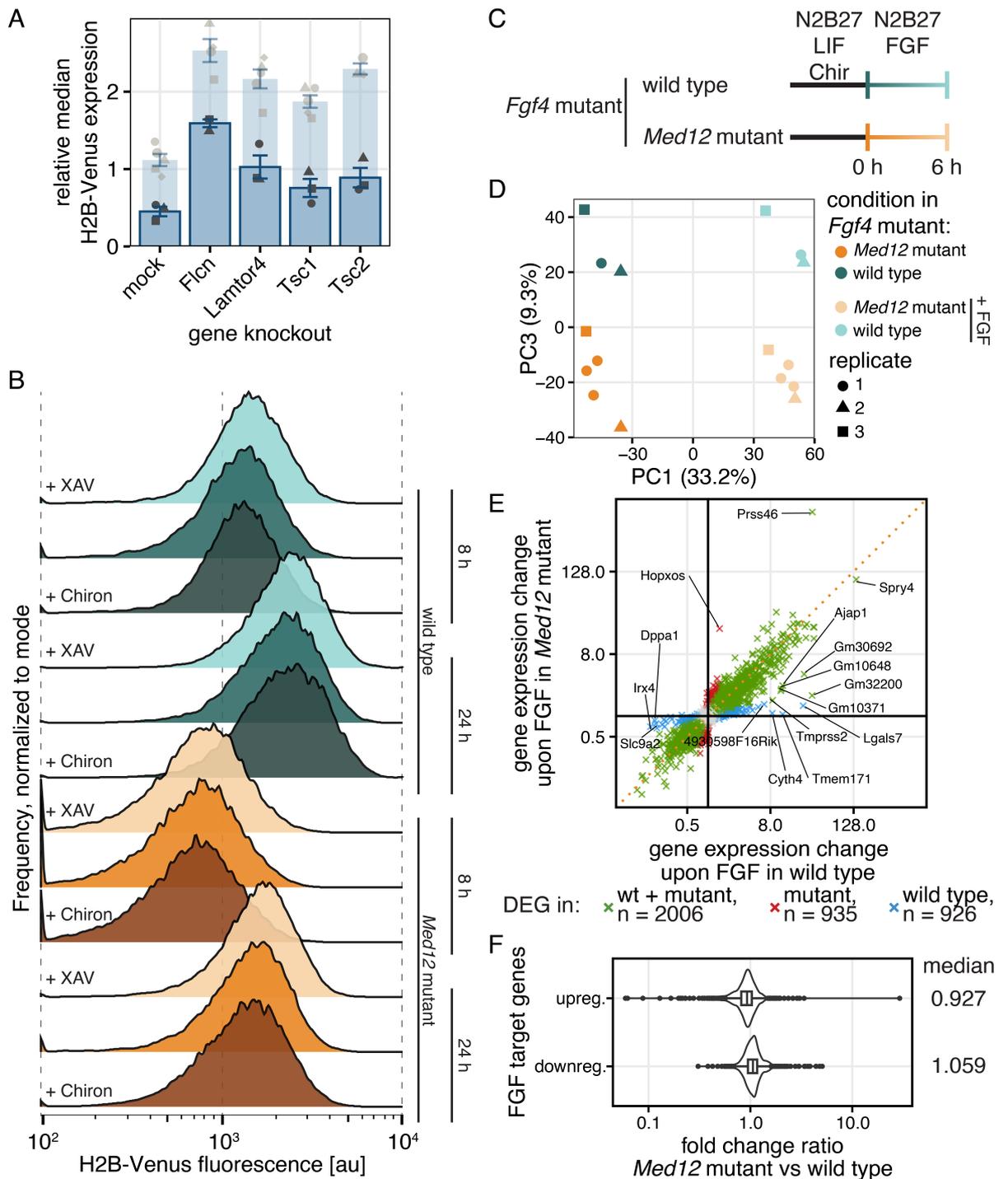


Figure 17: *Med12* regulates gene expression largely independent of mTOR, Wnt, and FGF-signaling.

A Median H2B-Venus fluorescence upon mutation of mTOR-related genes in *Med12*-mutant *Spry4*^{H2B-Venus/+} cells normalized to H2B-Venus expression in *Med12* wild-type cells. Median H2B-Venus fluorescence upon mutation of mTOR-related genes in *Med12* wild-type cells is reproduced from Figure 13D for comparison (light blue). Error bars indicate SEM, and points indicate individual replicates.

B *Spry4*^{H2B-Venus} expression levels in wild type (top) and *Med12* mutants (bottom) differentiated in N2B27 with different Wnt signaling environments for indicated time periods measured by flow cytometry.

C Schematic of an experiment to test *Med12*-dependency of FGF target genes by RNA sequencing.

D Principal component analysis transcriptomes from (F). PC2 (not shown; 12.8% of variance) separated experimental replicates from each other.

(Continued on next page)

E Gene expression changes, and the number of significantly differentially expressed genes (adjusted p-value < 0.01) upon FGF4 stimulation in wild-type versus *Med12*-mutant cells. The dotted orange line indicates the unity line. In addition to *Spry4*, the top 15 genes were selected for annotation based on their distance to the unity line.

F Ratio of fold changes for FGF target genes between wild-type and *Med12*-mutant cells, for up- (top) and downregulated genes (bottom). FGF target genes were defined as having a log₂-fold change in wild-type cells upon FGF4 stimulation > |1| and an adjusted p-value < 0.05. Bar indicates the median, and boxes indicate the 25th and 75th percentile.

4.3.4 Limited redundancy between *Med12* and *Med12L*

Since *Med12* had limited effects on the expression of signaling-induced gene expression changes, I asked if redundancy mitigates the effect sizes. The only known paralogue for *Med12* is *Med12L*, being 67% identical in the protein sequence with a similar domain structure (Luyties & Taatjes, 2022). Indeed, *Med12L* is upregulated in *Med12*-mutant lines, irrespective of the *Fgf4* genotype (Figure 18A,B). The fold change of upregulation on mRNA level varied between 2 and 3-fold. Interestingly, upon differentiation or stimulation with FGF4, *Med12L* expression decreased slightly, following the same trend in the wild-type and *Med12*-mutant lines. Polyclonal knockout of *Med12L* in *Med12*-mutant *Spry4*^{H2B-mVenus} cells in ESL + LIF revealed a further downregulation of the *Spry4*^{H2B-mVenus} reporter compared to the parental mock transfected *Med12*-mutant line. However, this difference was smaller than the effect of the loss of *Med12*, thereby indicating that *Med12L* is unable to replace *Med12* efficiently functionally.

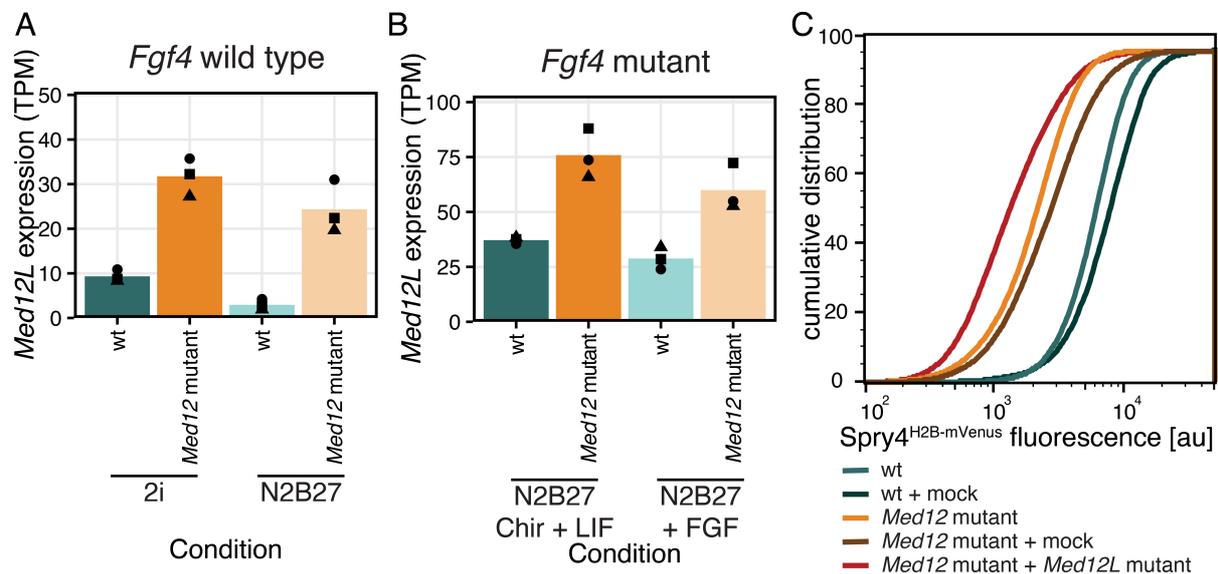


Figure 18: Limited redundancy between *Med12* and *Med12L*.

A,B Expression of *Med12L* in wild-type and *Med12*-mutant cells in a *Fgf4* wild-type (A) and *Fgf4*-mutant background (B). Data from experiments described in Figure 15A (A) and Figure 17C (B).

C H2B-Venus expression in untransfected wild-type and *Med12*-mutant cells and after transfection with mock or *Med12L*-targeting gRNAs. Expression was measured by flow cytometry 7 d after transfection and culture in ES + LIF medium.

4.4 *Med12* regulates pluripotency states and transitions

4.4.1 Lower clonogenicity in *Med12* mutants in the presence of FGF

Although MED12 was previously linked to the transcriptional regulation in response to specific pluripotency-related signaling systems (Lynch et al., 2020; Rocha et al., 2010), no strong dependence between responding transcriptionally to any tested signal and *Med12* was found. However, the decreased expression of the differentiation-associated gene *Spry4* in *Med12* mutants and increasing differences during differentiation between wild-type and *Med12* mutant cells suggested a role for *Med12* during early differentiation. To further describe the effects of the loss of *Med12* during the exit of pluripotency phenotypically, I performed a clonogenicity assay analyzing the ability of cells to react to changes in the signaling environment. Cells were differentiated in N2B27 for 2 d, reseeded at defined single-cell density into pluripotency medium, and the pluripotent colonies formed were stained and counted after 5 days (Figure 19A). The colony counts were normalized to the number of colonies formed from cells maintained in pluripotency medium all time. During the 2 d of differentiation in N2B27, wild-type cells did not lose pluripotency irretrievably. Compared to previously performed clonogenicity assays (Kalkan et al., 2017, 2019), the loss of pluripotency was delayed by the presence of LIF before the start of differentiation (Lackner et al., 2021). However, in *Med12*-mutant cells, fewer pluripotent colonies formed (Figure 19B). Surprisingly, even apart from the remaining pluripotent colonies, *Med12*-mutant cells grew with a limited proliferation speed, forming disorganized, flattened groups of cells with low alkaline phosphatase staining intensity (Figure 19B, right). Counting of pluripotent colonies revealed that *Med12* caused a reduction by one-third in colony number (Figure 19C). To investigate the relationship between the observed phenotype and the activity of specific signaling systems implicated in pluripotency exit, I conducted the same colony differentiation assay within an *Fgf4* mutant background. In this background, following the 48 h long differentiation period in N2B27 without FGF, both *Med12* wild-type and mutant cells exhibited a comparable colony formation rate upon transfer to pluripotency medium (Figure 19D, left). Supplementation of N2B27 with FGF4 revealed a drastic reduction in pluripotent colonies formed by *Fgf4; Med12* double mutant cells (Figure 19D, right). This reduction was significant in comparison to the *Fgf4* single mutant supplemented with FGF4, as well as the *Fgf4; Med12* double mutant in the absence of FGF4 (Figure 19E). Consequently, this experiment validates the premature commitment from naïve pluripotency upon *Med12* loss when FGF signaling is active. Further, it supports the conclusion from the transcriptomic analysis that *Med12* and FGF signaling work together during pluripotency exit.

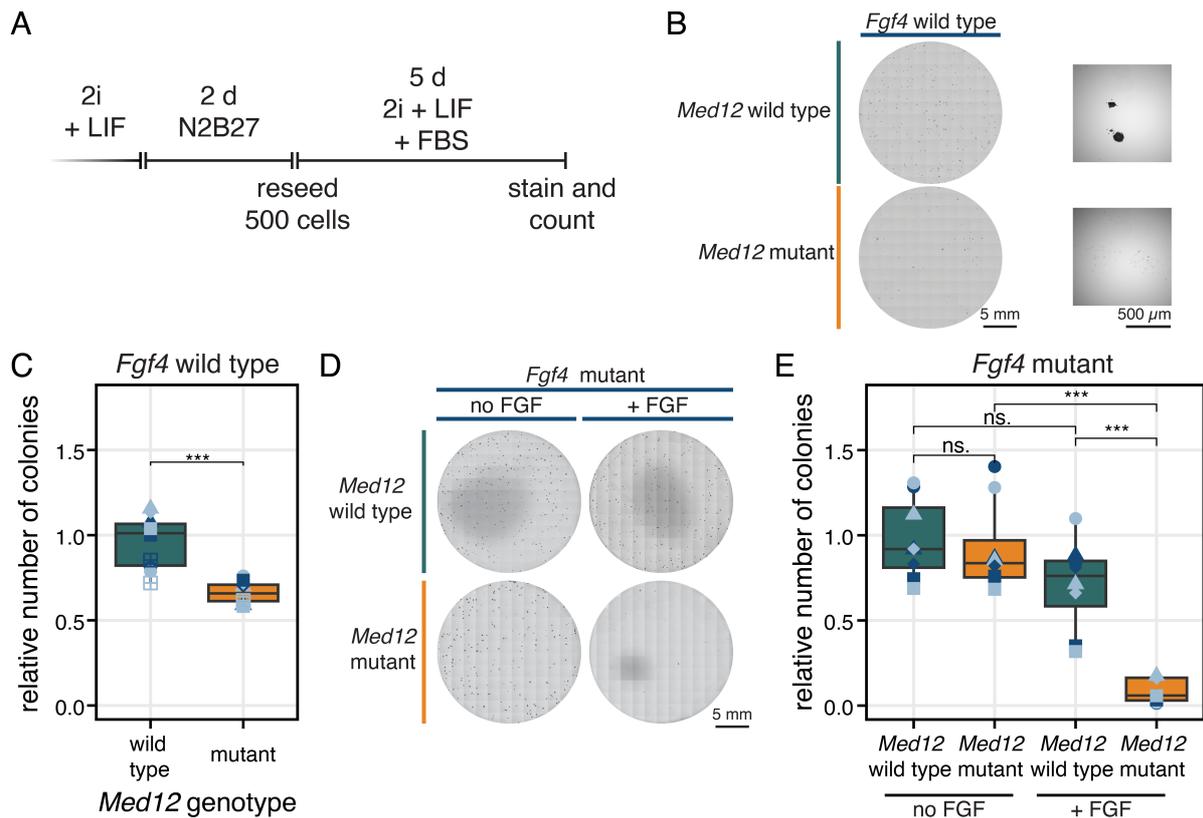


Figure 19: Mutation of *Med12* decreases clonogenicity.

A Experimental approach to determine clonogenicity of *Med12* mutants compared to wild-type cells.

B, C Representative images of dishes with formed colonies in the clonogenicity assay depicted in (A). (B) shows colonies formed in wild-type and *Med12*-mutant cells in an *Fgf4* wild-type background, (C) shows colonies formed in wild-type and *Med12*-mutant cells in an *Fgf4*-mutant background without (left) and with (right) FGF4 supplementation. (B, right) includes a close-up of typical pluripotent colonies (top) and cell outgrowth of *Med12* mutant cells (bottom).

D Number of colonies after treatment as indicated in (A) for both wild-type and *Med12*-mutant cells in an *Fgf4* wild-type (left) and *Fgf4* mutant background with and without supplementation with 10 ng/mL FGF4 (right). In (C) and (D): ns. indicates $p \geq 0.05$, *** indicates $p < 0.001$, paired Wilcoxon signed rank test. The bar indicates the median, and the boxes indicate the 25th and 75th percentiles.

4.4.2 *Med12* mutant cells leave naïve pluripotency delayed

The efficiency of how many colonies are formed during the clonogenicity assay depends on three factors: First, the starting conditions. If the exit of naïve pluripotency is considered a linear trajectory, affecting the starting conditions by *Med12* mutation might affect differentiation timing. Second, the gene expression changes in response to the changes in the extracellular signaling environment could be slower or delayed. And lastly, the formation of colonies after differentiation not only depends on the differentiation itself but also on the cell's ability to adapt again to changing signaling environments when transferring single cells back into the pluripotency medium. To understand how *Med12* loss of function led to lower clonogenicity in the exit of pluripotency assay, I reused the transcriptomic data comparing pluripotency and differentiation in N2B27 in wild-type and *Med12*-mutant cells (see Figure 15).

Exploiting a recently published (Lackner et al., 2021) time-resolved differentiation dataset, I quantified the relative difference in differentiation timing between wild-type and *Med12*-mutant cells as the Euclidean distance in the expression change in naïve pluripotency marker genes between data from this thesis and every timepoint of the time-resolved dataset. The wild-type cells from this study appeared to downregulate the pluripotency marker genes faster than in the transcriptomic dataset from literature (Figure 20A, top row), potentially due to different strains of mESCs (E14tg2 versus RC9). Relative to its parental wild-type line, the *Med12*-mutant cells differentiate slower, leading to a delay of 4 to 6 h after 24 h of differentiation (Figure 20A, bottom row). Contradictory, comparing the expression levels of the same set of naïve pluripotency marker genes directly between wild-type and *Med12*-mutant cells revealed that in cells kept in 2i all time, the naïve pluripotency markers are all slightly downregulated in the *Med12*-mutant cells (Figure 20B). During the 24 h of differentiation, naïve marker genes were downregulated in both genotypes, however, with a smaller slope of downregulation in *Med12*-mutant cells. Although the *Med12*-mutant cells were less pluripotent initially, these smaller expression changes led to a higher expression of pluripotency genes at the endpoint of the differentiation. This analysis was extended to an unbiased list of the 100 most strongly downregulated genes. Most genes showed a slope below 1, not strongly dependent on the genotype in which the genes were determined (Figure 20C). Similar trends were observed for upregulated genes (Figure 20D). These findings suggest that the delay quantified in Figure 20A is an underestimate since this quantification was not corrected for the different starting conditions, which position *Med12*-mutant cells as further differentiated. Overall, smaller gene expression changes in *Med12*-mutant cells, which additionally formed fewer colonies in the clonogenicity assay, suggest that reacting efficiently to reseed into the pluripotency medium is impaired in *Med12* mutants. Overall, this suggests a general role of *Med12* during the switch of transcriptional programs independent from any specific signaling systems which could lead to both phenotypes.

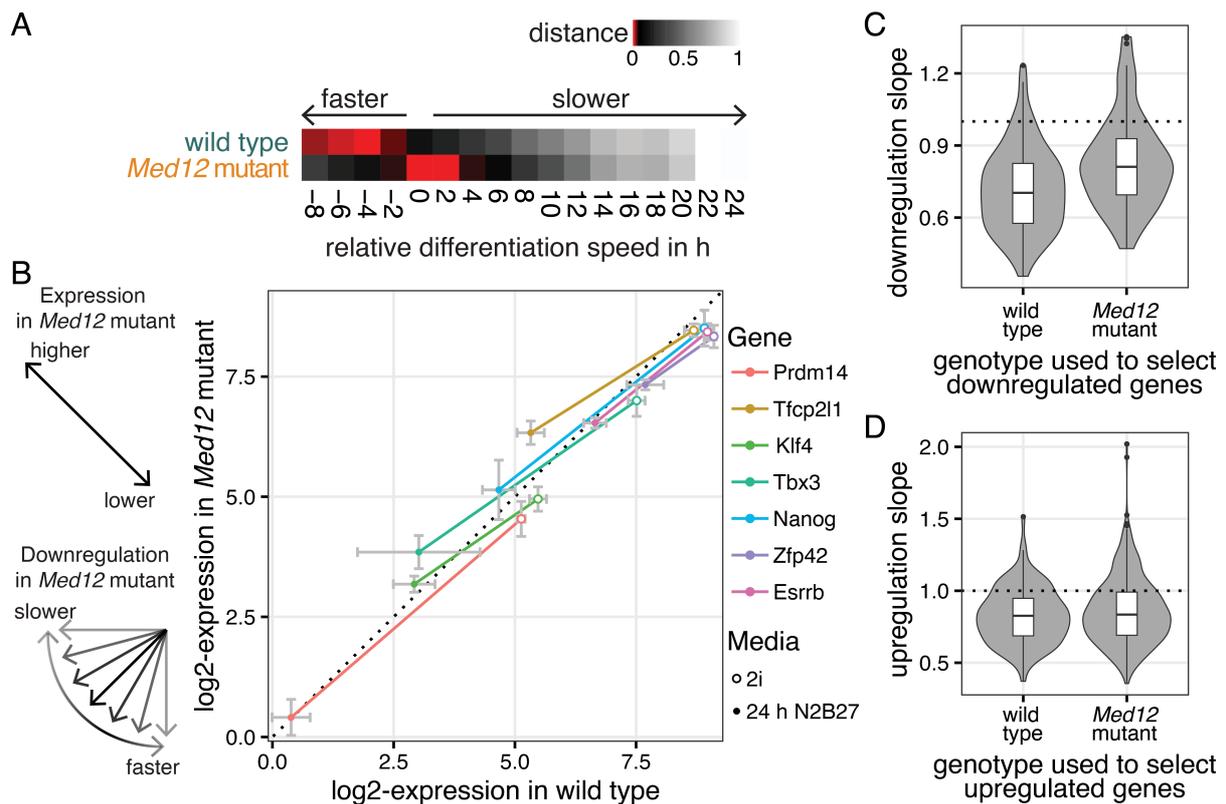


Figure 20: Differentiation delay in *Med12*-mutant cells.

A Estimation of differentiation delay in wild-type and *Med12*-mutant cells relative to a published time-resolved gene expression dataset (Lackner et al., 2021). Plot shows the normalized Euclidean distance of the expression of naive marker gene panel (*Prdm14*, *Tfcp2l1*, *Klf4*, *Tbx3*, *Nanog*, *Zfp42*, *Esrrb*) to the reference dataset. The negative delay values for wild-type cells from this study likely reflect small differences in experimental design compared to the study by Lackner et al., 2021.

B Expression of naive pluripotency marker genes in 2i (open circles) and after 24 h differentiation in N2B27 (dots) in wild-type versus *Med12*-mutant cells. Data from experiment in Figure 15. Relative expression values are shown as log₂-transformed TPM, and error bars indicate standard deviation.

C,D Distribution of downregulation (C) and upregulation (D) slopes determined as in B for the 100 genes with the strongest negative (C) and positive (D) fold-change in wild-type (left) and *Med12*-mutant cells (right). The bar indicates the median; the boxes indicate the 25th and 75th percentiles.

4.5 Establishing Epi, PrE, and AVE cells in *Med12*-mutant cells

4.5.1 Inducible GATA6 cell line to model PrE development

During exit of pluripotency, *Med12* enabled high efficiency of linear lineage transitions by implementing transcriptional changes. To further analyze how *Med12* is involved in establishing multiple cell types from a single precursor cell type, I next focused on the effects of a *Med12* mutation in epiblast (Epi) and primitive endoderm (PrE) differentiation from pluripotent cells. mESC cells can establish robust proportions of both cell types upon sufficient expression of GATA transcription factors, which can be accomplished by using an inducible transgene coding for GATA4 or GATA6 (Schröter et al., 2015). Epi and PrE differentiation are centered around a mutual inhibition of the transcription factors GATA and NANOG. By inducing GATA factors, both

transcription factors are temporarily co-expressed, and cell signaling via FGF establishes robust proportions of Epi cells, marked by NANOG, and PrE cells, characterized by upcoming endogenous GATA and SOX17 expression. Cells receiving FGF and GATA levels above a threshold level differentiate towards the PrE fate and inhibit their FGF secretion, balancing FGF levels in close surroundings (Figure 22A, Raina et al., 2021; Schröter et al., 2015). I established a GATA6-inducible line with similarly high induction levels as the previously used GATA4-inducible line (Figure 21A; Raina et al., 2021; Schumacher et al., 2023) to better mimic the order of induction events in the mouse embryo (Plusa et al., 2008). As reported previously (Schröter et al., 2015), GATA6 induction led to a similar capacity to differentiate into PrE and Epi cells, with slightly higher proportions of SOX17-positive PrE cells (Figure 21B,C).

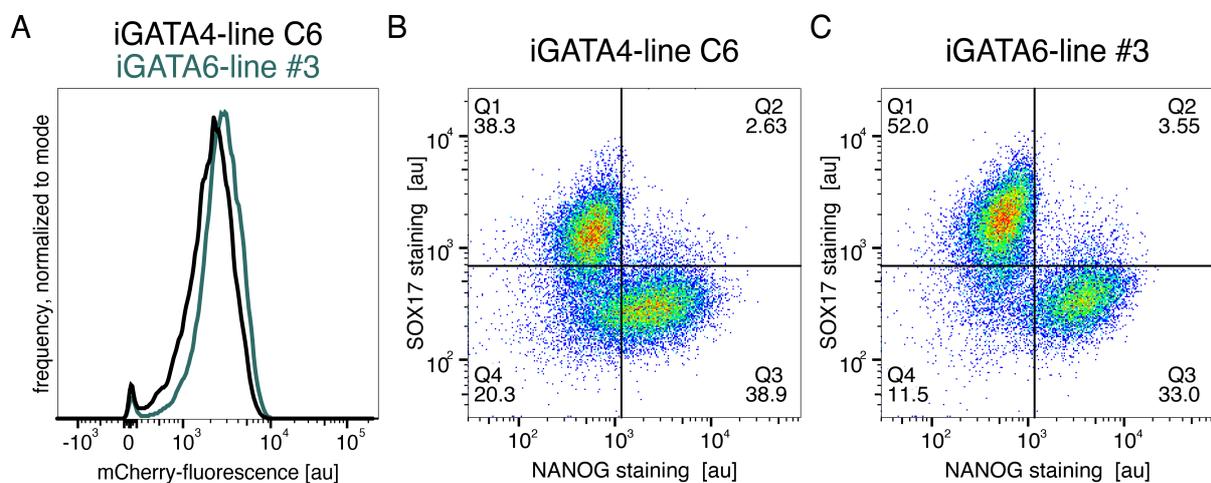


Figure 21: *Gata6*-inducible line with induction level and capability to differentiate into PrE similar to *Gata4*-inducible line.

A Induction levels of inducible GATA-mCherry transgene in previously established GATA4-line (black; Raina et al., 2021) and new GATA6-line (green) after 8 h of dox induction.

B,C Flow cytometry of differentiated cells (40 h in N2B27) into PrE and Epi from GATA4-inducible (B) and GATA6-inducible (C) mESCs. SOX17 marks PrE cells, NANOG marks Epi cells, and cell type proportions are indicated in plot corners.

4.5.2 Transitions to Epi and PrE identities buffered against loss of *Med12*

Using this GATA6-inducible cell line, I analyzed the effects during PrE differentiation in a *Med12*-mutant line (Figure 14A,B,E; Figure 22A). Using the previously established differentiation protocol, based on 8 h of doxycycline induction in the presence of 2i + LIF and differentiation in N2B27 for 20 h, a mixture of both cell types was established in both genotypes marked by mutual expression of NANOG (Epi) and SOX17 (PrE; Figure 22B). However, in the *Med12*-mutant line, fewer cells were exclusively SOX17-positive, with no major differences in the abundance of double positive or negative cells (Figure 22C). The proportion of cells differentiating to PrE increased with supplementation of FGF during the 20 h differentiation period in both wild-type and *Med12*-

mutant cells. Still, the difference between the genotypes was maintained with 93.1% PrE cells in wild-type versus 51.4% in *Med12*-mutant cells (Figure 22B,C). This suggests that lower FGF/ERK signaling levels do not cause the lower PrE proportions in the *Med12*-mutant cells, in line with similar ppERK levels between wild-type and *Med12*-mutant cells in ES + LIF (Figure 14E). The efficient differentiation of mESCs towards PrE cells depends on expressing the transcription factor GATA at levels beyond a threshold level. Given this dependency, I analyzed the expression levels just after stopping the induction of GATA6 with doxycycline. Independent from the integration sites of the GATA transgene and transgene induction levels in the parental lines, the loss of *Med12* caused lower induction levels in multiple clonal iGATA6 lines (Figure 22D). In wild-type cells, the proportion of the two cell types is not dependent on GATA induction levels within a specific range (Schröter et al., 2015). Still, the lower proportions in the *Med12*-mutant cells could result from cells not reaching the threshold of GATA induction levels, thereby never having the potential to differentiate into PrE. To correct for this possibility, I aligned the GATA induction levels after induction for 8 h by flow sorting both genotypes for a specific GATA6-mCherry induction level range. The sorting gate defined this range for approximately 25% of the *Med12*-mutant cells with as high induction levels around the average induction levels in wild-type cells. Cells from both genotype passing this gate were reseeded for differentiation (Figure 22E, left). As expected, this led to aligned GATA6-mCherry levels immediately after sorting (Figure 22E, right). However, sorting and reseeded the cells disrupted the local FGF signaling environment, causing very low PrE differentiation efficiency in both wild-type and *Med12*-mutant cells in N2B27 (Figure 22F,G). The addition of exogenous FGF overcame this limitation, leading to very high PrE differentiation efficiency in presence of FGF4 (Figure 22F,G). This suggested that all *Med12*-mutant cells were still able to establish a PrE fate, but reduced GATA6-mCherry induction levels limit the proportion of cells with the potential to do so. Taking into account that highly induced *Med12* mutant cells fully converted to PrE cells in the presence of FGF, these experiments suggested that while *Med12* plays crucial roles in regulating the expression of specific genes, the transition from embryonic to extraembryonic identities is largely buffered to the loss of *Med12*.

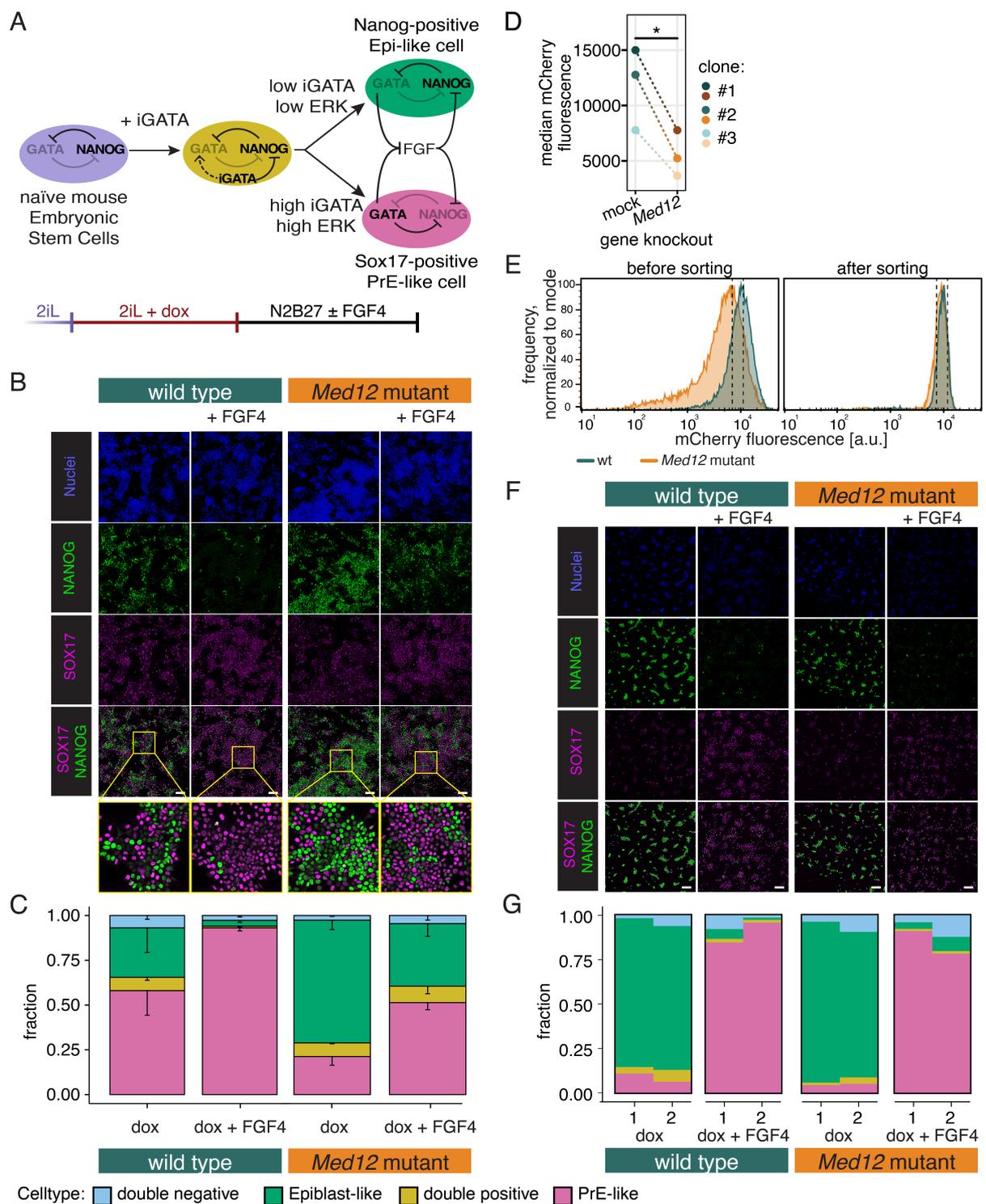


Figure 22: Transition between embryonic and extraembryonic identities is buffered against loss of *Med12*.

A Schematic of experimental approach to model differentiation of mESCs towards epiblast and primitive endoderm via GATA induction.

B Immunostaining for the Epi-marker NANOG (green) and the PrE marker SOX17 (magenta) after 8 h of GATA6 induction followed by 20 h of differentiation in wild-type and *Med12*-mutant cells, including a closeup of the merged channel in the lowest panel. Scale bars: 100 μ m.

C Quantification of cell type proportions after differentiating wild-type and *Med12*-mutant cells as in (B). N=3, n > 1100 segmented cells per replicate.

(Continued on next page)

D Median GATA6-mCherry fluorescence upon 8h dox induction in three independent clonal GATA6-mCherry inducible cell lines 7 days after transfection with control or *Med12*-targeting gRNAs. * indicates $p < 0.05$, paired student's t-test.

E GATA6-mCherry fluorescence after 8 h of dox induction. The left shows the distribution of expression levels in the whole population; right shows expression levels after flow sorting of cells with similar fluorescence intensity. Dashed lines indicate the sorting gate.

F Immunostaining of the Epi-marker NANOG (green) and the PrE marker SOX17 (magenta) after 8 h of GATA6 induction, flow sorting as described in (E), reseeded, and 20 h of differentiation with and without exogenous FGF4 in wild-type and *Med12*-mutant cells. Scale bar: 100 μm .

G Quantification of cell type proportions after differentiating wild-type and *Med12*-mutant cells as in (F). $N=2$, $n > 500$ cells per replicate.

To further corroborate the dependence of *Med12*-mutant cells on GATA6 induction levels, I aimed to align the GATA6 induction levels at the end of induction by delaying the start of induction in wild-type cells by 4 h. To follow GATA6 induction levels in single cells over time and track nuclei, the GATA6-inducible line was tagged with a stably expressed H2B-Cerulean by transduction (Figure 23A, left). Additionally, cells were stained at the end of the time-lapse acquisition and differentiated in FGF supplemented N2B27 for the cell type markers SOX17 and NANOG to link tracked cells to their final differentiation result (Figure 23A, right). Quantifying the nuclear GATA6-mCherry fluorescence showed differences between wild-type and *Med12*-mutant cells even with the adapted doxycycline induction conditions. In the wild-type, GATA6-mCherry levels rose faster and did not reach saturation within the 4 h of induction (Figure 23B). Highest GATA6-mCherry levels were detected only after the end of induction. In contrast, *Med12*-mutant cells expressed GATA6-mCherry at levels approaching saturation at the end of the induction period, with a lower maximal expression than wild-type cells. Despite the maximal expression of GATA6-mCherry was higher in wild-type than in *Med12*-mutant cells, the overall abundance of the transgene, measured by the integrated fluorescence intensity over time, was not significantly different (Figure 23C). Single cells were tracked over the entire time series and annotated into PrE and Epi cells (Figure 23D). At 10 h after the start of the time series, corresponding to 2 h after the end of induction, PrE cells had slightly higher GATA6-mCherry expression levels than Epi cells from the same timepoint in both wild-type and *Med12*-mutant cells (Figure 23E). To further analyze the dependence of differentiation outcome on *Gata6* induction levels, I applied receiver operator characteristics and calculated the area under the curve (AUC) as a measure of predictability of the optimal threshold best separating PrE and Epi cells at each time point (Fawcett, 2006). GATA6-mCherry levels became a predictor of differentiation outcome early during its upregulation, reaching similar accuracy levels of 0.7 in wild-type and *Med12*-mutant cells (Figure 23F). The optimal threshold to separate PrE from Epi cells remained lower in *Med12*-mutant cells, reflecting the induction levels over time (Figure 23G). This analysis confirmed the dependency of the PrE differentiation on GATA6 induction levels on a cellular level for both

genotypes and explained the altered cell type proportions in the *Med12* mutants by its effect on GATA6 induction levels and dynamics.

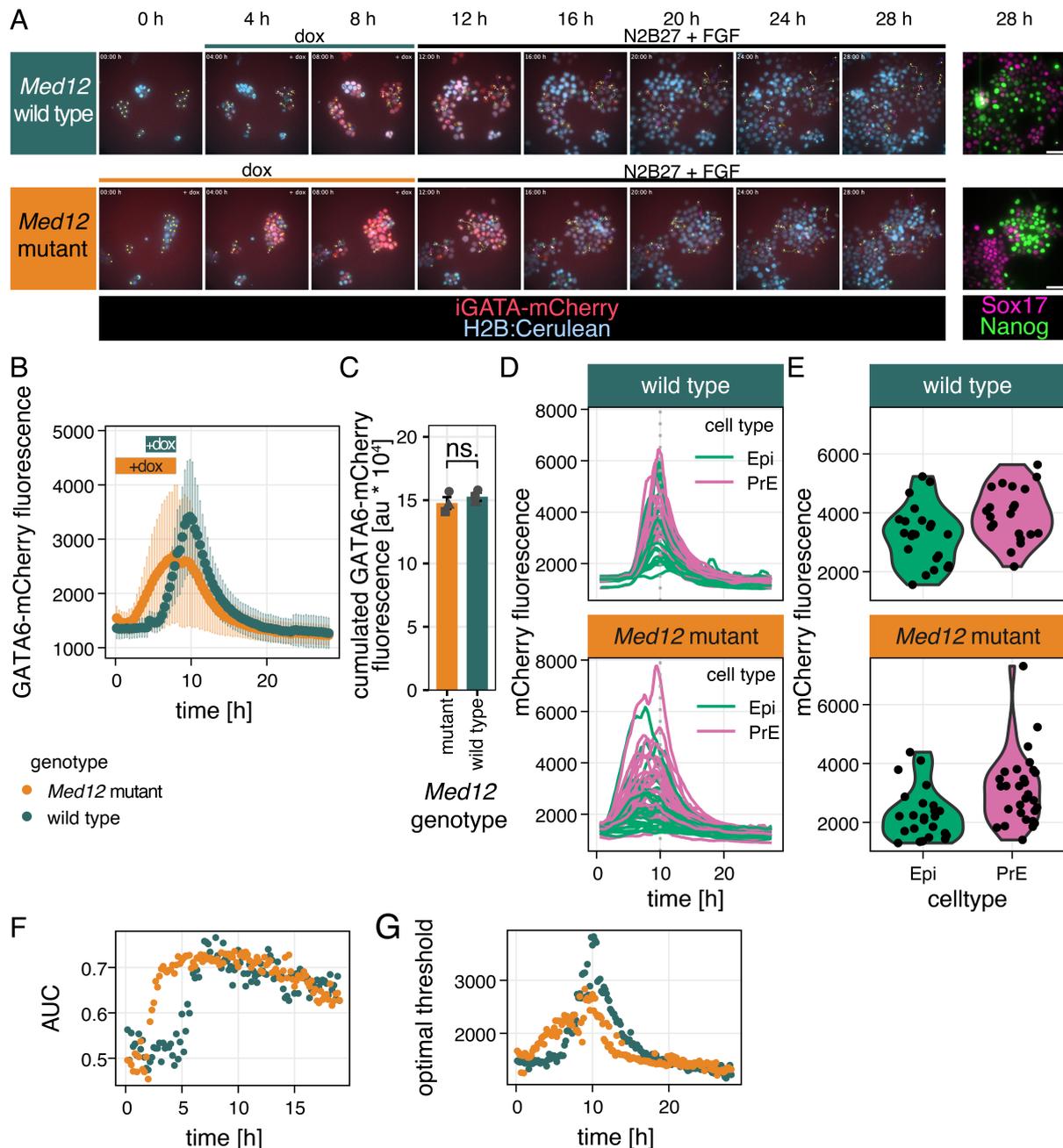


Figure 23: Loss of *Med12* affects Gata6-induction dynamics.

A Example stills of the time-lapse imaging during induction and differentiation of wild-type and *Med12*-mutant cells with H2B-Cerulean marked nuclei (red) and GATA6-mCherry levels (red). On the left: Immunostaining of the time-lapse imaged cells for SOX17 (marking PrE) and NANOG (marking Epi). Doxycycline treatment is indicated for each genotype.

B Quantification of GATA6-mCherry expression dynamics from time-lapse movies during induction and differentiation. Boxes indicate induction times (8 h for *Med12* mutant, 4 h for wild type). Error bars indicate SD. One out of N = 3 replicates shown, n > 400 cells per time point.

C Mean GATA6-mCherry fluorescence integrated over the total time-lapse imaging experiment from (B) in wild-type and *Med12*-mutant cells. Points refer to individual replicates, error bars indicate SEM. ns. Indicates p > 0.05, Wilcoxon signed-rank test.

(Continued on next page)

D Same experiment as in (B), but showing GATA6-mCherry fluorescence in single cells. Trace color indicates differentiation outcome determined by immunostaining (Epi: green; PrE: magenta).

E GATA6-mCherry fluorescence in single cells 2 h after the end of induction.

F Predictive power of GATA6-mCherry expression determined as Area Under the Curve (AUC) from ROC-analysis.

G Optimal GATA6-mCherry threshold to predict differentiation outcome determined by Youden's J statistic (F) of ROC-analysis.

D - G Data from one out of N = 3 replicates, n ≥ 45 cells per genotype.

4.5.3 *Med12* mutant cells can acquire AVE fate

Upon GATA-induction *Med12* mutant cells a significant proportion still adapted an PrE fate. I next sought to analyze if these *Med12*-mutant PrE cells were functionally equivalent to their wild-type correspondent. In the mouse embryo, a subpopulation of PrE-derived VE cells differentiates subsequently into AVE cells, marking the establishment of the anterior-to-posterior axis. *Med12*-mutant embryos failed to establish an AVE fate (Rocha et al., 2010). To challenge this, I aimed to differentiate *Med12*-mutant AVE cells in vitro. Recently, we established 2D in vitro differentiation conditions for AVE, first differentiating mESCs into a VE layer with extensive GATA induction and FGF supplementation, followed by activating Nodal signaling with Activin, still in the presence of FGF (Figure 18A, Schumacher et al., 2023). Using this protocol with the GATA6-inducible wild-type line confirmed the ability of this line to form AVE-like cells, defined by the expression of the marker genes *Eomes* and *Otx2* (Figure 24B). While *Eomes* is expected to localize to AVE regions, *Otx2* is additionally expressed in Epi-like cells (Schumacher et al., 2023). Epi-like identity was therefore assessed via POU5F1 expression. In both wild-type and *Med12*-mutant, non-differentiated cells without GATA-induction only expressed the naïve marker POU5F1 (Figure 24B, lower panel). While POU5F1 was absent upon the AVE differentiation protocol in the vast majority of wild-type cells, patches of *Med12*-mutant cells maintained its expression (Figure 24B, upper panel). This is most likely caused by the incomplete transition into a PrE state even upon GATA6 induction and FGF treatment (Figure 22). This was further confirmed by the mutually exclusive expression of the PrE marker GATA6 and the Epi marker POU5F1 (Figure 24C). However, within regions of successful PrE differentiation indicated by the absence of POU5F1, the AVE markers OTX2 and EOMES were expressed similarly to the wild type (Figure 24B). The potential of *Med12*-mutant cells to adapt an AVE-identity further proves the similarity to wild-type cells concerning the functionality of the PrE-cells.

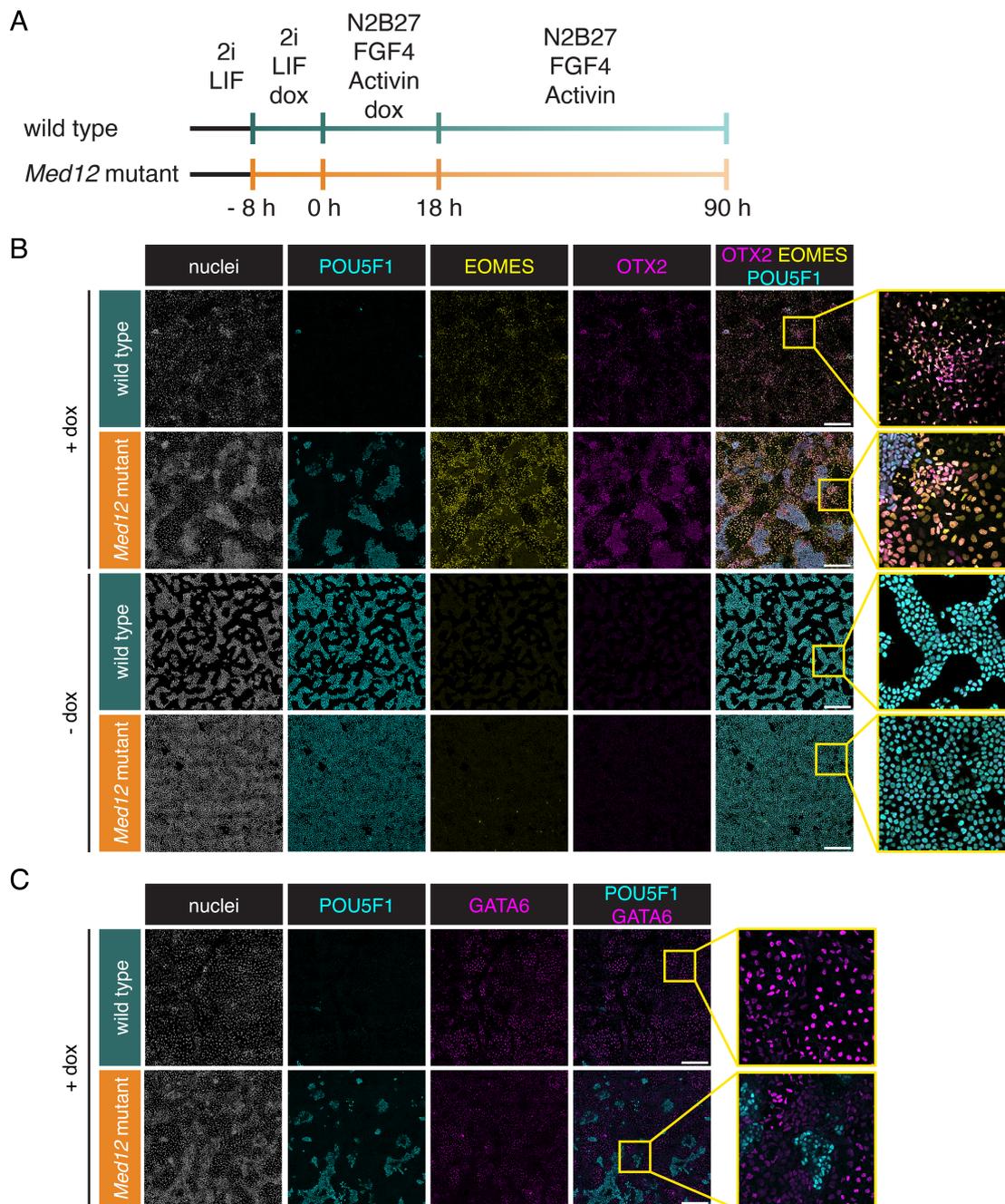


Figure 24: AVE differentiation in *Med12*-mutant cells.

A Schematic of the 2D AVE differentiation protocol established by Schumacher et al., 2023 used to compare wild-type and *Med12*-mutant cells. Time distances are not drawn to scale.

B Immunostaining for POU5F1 (cyan), EOMES (yellow), and OTX2 (magenta) in wild-type and *Med12*-mutant cells after treatment as described in (A) as well as a control without doxycycline induction.

C Immunostaining for POU5F1 (cyan) and GATA6 (magenta) in wild-type and *Med12*-mutant cells after treatment as described in (A).

4.6 *Med12* enhances plasticity during PrE differentiation

4.6.1 Transcriptome-wide comparison of wild type and *Med12* mutant in PrE differentiation

Focusing on the marker gene expression NANOG and SOX17, *Med12*-mutant cells were able to form patterns of mutual exclusive expression, indicating formation of Epi and PrE identities. However, the expression of the GATA6 transgene was less efficient and previous analysis of *Med12*-mutant cells revealed decreased expression of *Spry4* and phenotypic effects on cell type transitions. To understand how *Med12* influences the transcriptional regulation during PrE differentiation transcriptome-wide, I performed a single-cell transcriptomics experiment, comparing wild-type and *Med12*-mutant cells. Single-cell resolution allowed the assessment of transcriptional heterogeneity within cell states. This approach will enable the detection of more subtle effects on lineage bifurcations upon loss of *Med12*. Two wild-type cell lines with GATA4 or GATA6 inducible gene constructs and three different clonal lines of the *Med12* mutant in the GATA6-inducible line were included as replicates. To minimize batch effects, all conditions were multiplexed enabled by labeled with cell multiplexing oligos. For droplet-based single-cell RNA sequencing the samples were combined in order to maximize sensitivity for transcriptomic differences between the genotypes and minimize batch effects (Figure 25A). Sequencing the cell multiplexing library together with the gene expression library enabled demultiplexing with robust separation between signal from cells and background (Supplementary Figure 7). Downstream filtering for high-quality cells resulted in between 1082 and 1763 cells per sample (Supplementary Figure 8, Supplementary Table 3). In UMAP space (McInnes et al., 2018), the 10 samples fell into two major groups, separated pluripotency from differentiated conditions (Figure 25B). The wild-type single-cell transcriptomes within those groups localized in distinct regions from *Med12*-mutant cells. While the clonal replicates of the *Med12*-mutant cells were well mixed along UMAP1 and UMAP2, the separation between the GATA4 and GATA6-inducible lines indicate transcriptomic differences between the lines. The separation between pluripotent and differentiated cells was further confirmed by the expression of the pluripotency-associated Wnt/ β -Catenin target gene *Sp5* (Figure 25C). *Nanog* and *Fgf4* expression were detected in pluripotent cells and a subgroup of differentiated cells, while *Dnmt3l* was only observed in differentiated samples. The PrE markers *Sox17*, *Dab2*, and *Cubn* were expressed in a subset of differentiated cells, complementing the *Nanog*-positive cells of the differentiated samples. These discrete regions of Epi and PrE marker genes indicate the split into two lineages in response to dox treatment and differentiation.

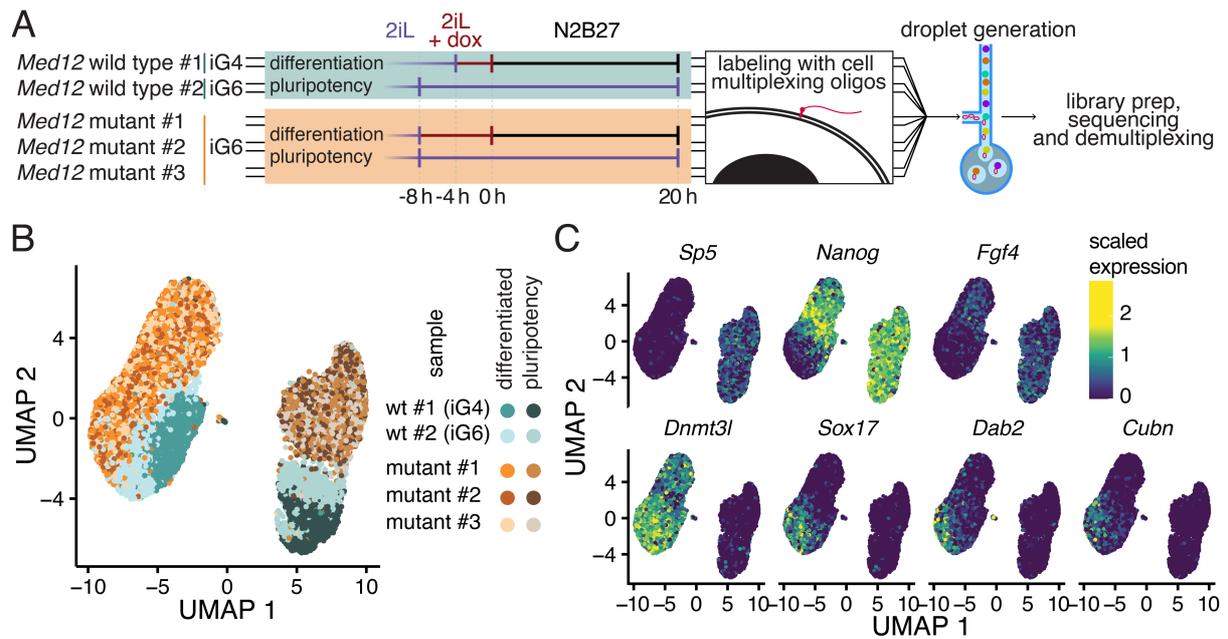


Figure 25: Multiplexed single-cell RNA sequencing of wild-type and *Med12*-mutant cells.

A Schematic of the single-cell RNA sequencing experiment to compare single-cell transcriptional signatures between wild-type (wt) and *Med12*-mutant cells in pluripotency and upon PrE differentiation. Doxycycline induction was 4 h and 8 h in wild-type and *Med12*-mutant cells, respectively. Cells were labeled with lipid-bound multiplexing oligos before pooled droplet generation, library preparation, sequencing, and computationally demultiplexing into separate samples.

B UMAP representation of single-cell transcriptomes from all samples.

C Ln-scaled marker gene expression projected on UMAP representation from (B).

4.6.2 *Med12* affects pluripotency homogenously

To survey the effects of the loss of MED12 in detail, I first focused on the single-cell transcriptomes in pluripotency. Principal component analysis showed that the two primary sources of variation between the cells depended on their *Med12* genotype (Figure 26A). In PC1 and PC2, explaining 28.1% of the total variance in the dataset, there was very limited overlap between wild-type and *Med12* mutant cells. Again, the three clonal *Med12*-mutant lines were well mixed, whereas a better separation between the two different GATA-factor inducible lines suggests small differences between *Gata4* and *Gata6*. These differences might not only reflect differences in the biological function of *Gata4* versus *Gata6* but also could reflect the consequences of the random integration of the transgene into the genome or different passage numbers. Markers of naïve pluripotency were downregulated (*Klf4*, *Zfp42*, *Tbx3*) or equally (*Prdm14*, *Nanog*, *Esrrb*, *Tcf2111*) expressed in *Med12*-mutant cells (Figure 26B), mainly reflecting the differences between those genotypes observed in the bulk RNA sequencing experiment (Figure 15). Visualization with non-linear dimensionality reduction using UMAP together with Louvain clustering (Nguyen et al., 2008) confirmed differences between the genotypes but further suggested more heterogeneity in the dataset, resulting in three clusters (Figure 26C). Most cells clustered within clusters 0 and 1,

which robustly separated *Med12* genotypes (Figure 26D). The third cluster only made up 1.1% of the total cell number. These cells likely corresponded to a few cells expressing the doxycycline-inducible transgene even without adding doxycycline to the medium since *Gata6* was specifically upregulated in these cells (Figure 26E). To understand the remaining heterogeneity between the single-cell transcriptomes and if this corresponds to any subsets of cells with different identities, I summarized the expression of genes previously described to have cell-cycle dependent expression rates into a relative score for each cell and cell-cycle phase. Groups of cells appearing closer together in the UMAP plot could be annotated to specific cell cycle phases by their cell cycle phase scores (Figure 26F). The local structures in the UMAP plot, are therefore likely to reflect cell cycle states of the corresponding single cell transcriptomes and resulted in similar groups and cell cycle phases for wild-type and *Med12*-mutant cells. The total number of mRNA counts was dominated by cell-to-cell variability in the local neighborhood and did not strongly reflect the cell size variations across the cell cycle (Figure 26G). In summary, in pluripotency, the loss of *Med12* resulted in a homogenous shift in their gene expression change, irrespective of the cell cycle as the particular major source for gene expression state heterogeneity.

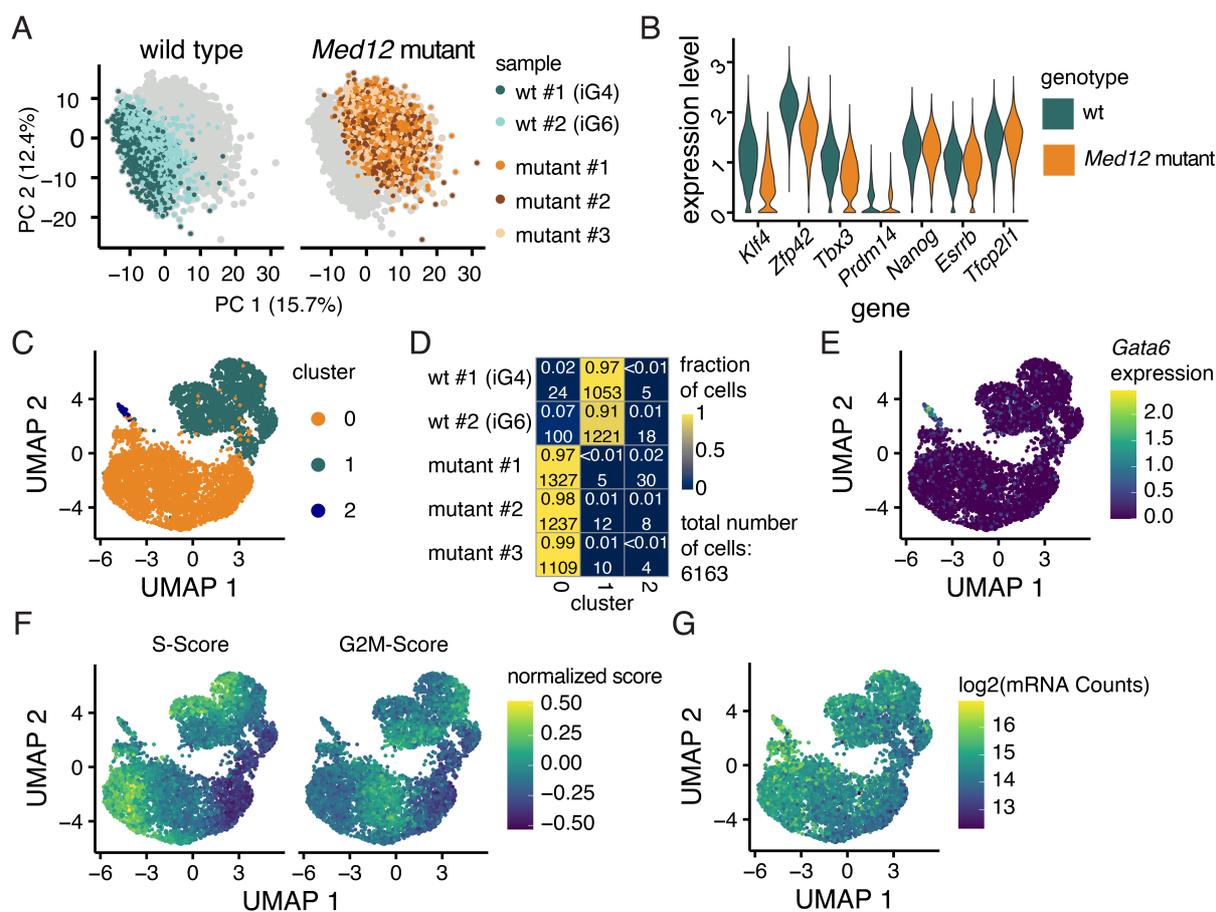


Figure 26: Loss of *Med12* affects pluripotency homogeneously.

A Principal component analysis separates both wild-type lines from *Med12*-mutant cells in pluripotency.

B Ln-transformed expression levels of same naïve pluripotency marker genes as in Fig. 4C,D.

(Continued on next page)

C Clustering of single-cell transcriptomes from (A), shown in UMAP space.

D Heatmap showing the fraction and total number of cells from each sample in the three clusters from (C).

E Ln-transformed expression levels of *Gata6* in the same UMAP plot as in (C).

F Normalized cell cycle score for S-phase and the G2M transition in the same UMAP plot as in (C).

G Total detected mRNAs per cell in the same UMAP plot as in (C).

4.6.3 Lower mRNA counts in *Med12*-mutant cells upon differentiation

Med12 was identified as a general transcriptional regulator in mESCs. However, loss of MED12 affected signaling-induced exit of pluripotency, while the establishment of Epi and PrE cell types was largely buffered against loss of *Med12*. To investigate how *Med12* affects transcription during cell type transitions, the mRNA abundance per cell was compared between wild-type and *Med12*-mutant cells. Due to the multiplexed design of the single-cell RNA sequencing experiment, the total number of detected mRNAs is expected to correlate with the total mRNA abundance since no systematic batch effects in sample lysis, library preparation, and sequencing depth occur. In pluripotency conditions, the number of mRNAs detected per cell was not significantly different between wild-type and *Med12* mutant cells (median number of UMIs 24437 and 23806, respectively, Figure 27A, Supplementary Table 3). Differentiation resulted in an increase in mRNA abundance in both genotypes (Figure 27A), a general trend observed previously comparing 2i + LIF and ES + LIF medium (Shao et al., 2022). In wild-type cells, the detected mRNAs increased by 15.5% during differentiation; in *Med12*-mutant cells, this increase was with 7.7% only half as in the wild type (Figure 27A). Plotting total number of detected mRNAs for each cell in the UMAP dimensionality reduction confirmed that this increase in mRNA abundance is not due to a transcriptionally distinct outgroup but rather a general shift in the distribution of mRNA counts per cell (Figure 27B). These analyses of mRNA abundance in wild-type and *Med12*-mutant cells suggest that the presence of MED12 amplifies the upregulation of global transcriptional output during differentiation.

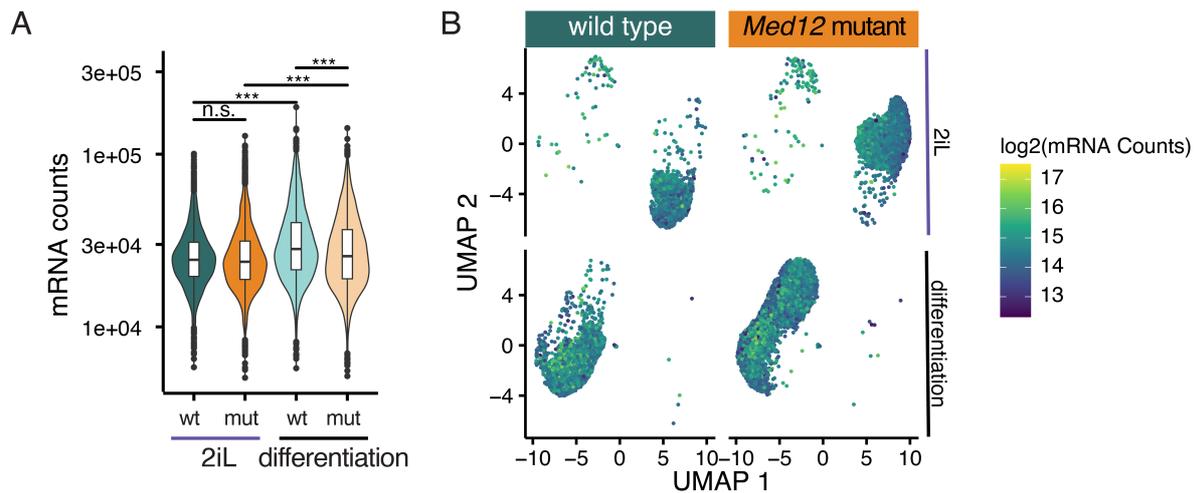


Figure 27: mRNA counts increase less during differentiation in *Med12*-mutant cells.

A Number of detected mRNAs per cell in wild-type and *Med12*-mutant cells. ns. indicates $p \geq 0.05$, *** indicates $p < 0.001$, Bonferroni-adjusted Kolmogorov-Smirnov test.

B Log₂-transformed mRNA counts for every cell in UMAP space, split by *Med12* genotype and media condition.

4.6.4 *Med12* has only minor effects on cell identities

Lower transcriptional output in *Med12*-mutant cells during differentiation suggested an influence of *Med12* on differentiation outcome. The marker gene expression of *SOX17* and *NANOG* only revealed differences in cell type proportions due to lower induction levels, but in general, differentiation into Epi and PrE appeared to be buffered against the loss of *Med12*. To further analyze if the loss of *Med12* further influenced the established cell identities, I integrated the single-cell transcriptomes of the differentiated samples with each other. This approach equalized differences dependent on the sample and, thereby, genotype for each cell type, but did not influence the difference between the cell types within the samples. Principal component analysis, combined with Louvain clustering, led to colocalization between the samples of two major clusters, separated mainly in PC1 (Figure 28A). Based on the identity of differentially expressed genes, most prominently *Gata6* and *Nanog* (Figure 28B), these clusters corresponded to PrE and Epi identities. The influence of *Med12* on the proportions of the two cell types was confirmed by the unbiased clustering (Figure 28C). This clustering resulted in slightly higher proportions of PrE cells (between 35% and 42% for *Med12* mutants, 60% and 77% for wild-type lines, Figure 28C) than the definition of cell types based on the marker genes *Sox17* and *Nanog* in immunostainings (21% in *Med12* mutants, 58% in wild type, Figure 22C). With these annotations to PrE- and Epi-cells, I compared gene expression changes during differentiation in wild-type and *Med12*-mutant cells. Genes having a log₂ gene expression fold change > 0.5 in wild-type cells between pluripotency and differentiation showed a unimodal distribution of fold changes in *Med12*-mutant cells (Figure 28D). This was confirmed for downregulated and upregulated genes and

differentiation into both cell types. Thus, *Med12* does not seem to be required to express specific modules of genes among all differentiation-associated genes. This conclusion was strengthened further by globally analyzing all expressed genes and their expression change during Epi (Figure 28E) and PrE (Figure 28F) differentiation. No class of genes appeared to be specifically differentially expressed only in wild-type cells, similar to the absence of a subset of FGF target genes specifically dependent on *Med12* (Figure 17). Notably, effects only reliant on the differences in total mRNA counts per cell between the two genotypes would not be detected here due to the normalization to total counts per cell for detecting differentially expressed genes.

Despite *Med12* not being essential for regulating Epi and PrE-specific genes, differentially expressed genes between the genotypes were identified in all three states: Pluripotency, Epi, and PrE. Comparing the top downregulated genes in these lists revealed that their identity largely overlaps for all three cell states (Figure 28, left), while the overlap was smaller for upregulated genes (Figure 28, right). Genes that were specific for one of the states often mark genes that are only expressed in this particular state, e.g., *Klf4* and *Zfp42* in pluripotency. In conclusion, rather than being cell differentiation-specific, differences between wild-type and *Med12*-mutant cells were shared between different cell identities.

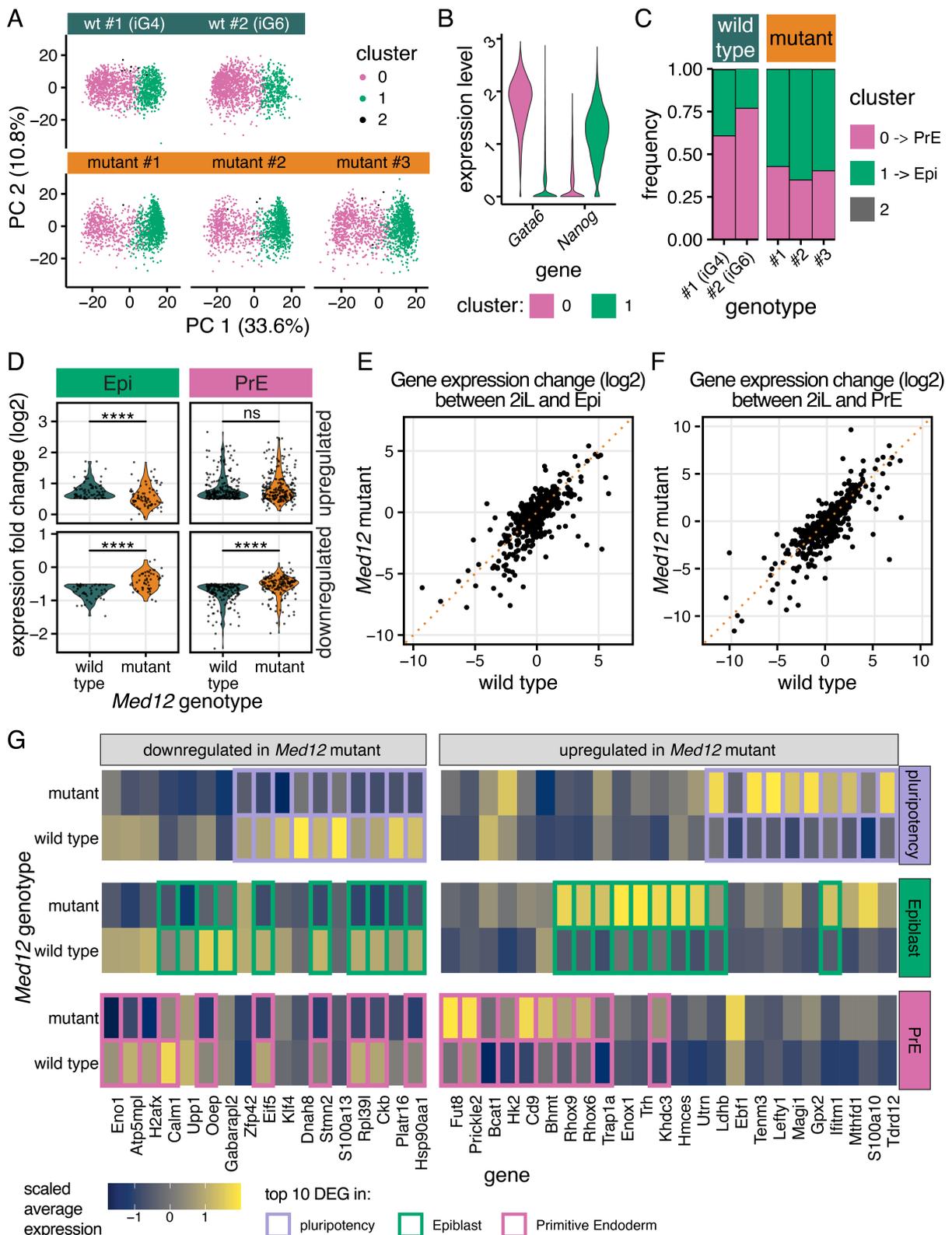


Figure 28: Limited role of *Med12* in cell type-specific gene regulation.

A Principal component analysis and Louvain-clustering of single-cell transcriptomes from wild-type and *Med12*-mutant cells after differentiation.

B Ln-transformed expression levels of the PrE-marker gene *Gata6* (magenta) and the Epi-marker gene *Nanog* (green) in the differentiated samples split by cluster. Cluster 2 was excluded due to the small number of cells.

(Continued on next page)

C Proportions of cell types, identified by clustering, in wild-type and *Med12*-mutant cells.

D Comparison of up- (top panels) and downregulated genes (bottom panels) in wild-type and *Med12*-mutant cells upon differentiation from pluripotency to Epi (left) or PrE (right). Shown are genes with a log₂-change of expression > 0.5 in wild-type cells. ns. indicates $p \geq 0.05$, **** indicates $p < 0.0001$, paired Wilcoxon signed rank test.

E and F Expression change of each gene upon differentiation from pluripotency (2i + LIF) to Epi (E) and PrE (F) in wild-type versus *Med12*-mutant cells. The dotted orange line indicates the unity line.

G Differentially expressed genes between *Med12* wild type and mutant cells for the three cell states. Tile color shows scaled average gene expression; colored boxes indicate the 10 genes with the largest fold-change between *Med12* wild-type and mutant cells in each cell state.

4.6.5 *Med12* mutants are less plastic and have lower biological noise

The effect of *Med12* on lineage-specific gene expression was minor, however, the impact on general transcriptional efficiency might still influence the formation of Epi and PrE cells. In the PCA analysis of the differentiated samples, the two cell types appeared to be less separated along PC1 (Figure 28A). To quantify how well the clusters defining the cell types are separated, I calculated the Jaccard similarity index of 100 iterations of clustering with a random subset of 70% of the cells compared to clustering of the total dataset per sample. This measure was previously established to assess cluster stability (Tang et al., 2021). In both genotypes, the Jaccard index reached > 0.85 in all iterations for both genotypes, indicating robust clustering (Figure 29A). Nevertheless, wild-type cells clustered significantly less robustly compared to the *Med12*-mutant cells, consistent between cell or clonal lines (Figure 29A). The absence of *Med12* resulted in an improved transcriptional segregation of the cell types. Moreover, the extended tail of the distribution in the wild type implies the existence of a larger cellular population exploring transcriptional space in-between the primary clusters corresponding to Epi and PrE cells. The lack of populating this intermediate region between the cell types could be dependent on differences in the transgene induction dynamics, or point towards a reduced capacity of *Med12* mutant cells to switch between differentiation trajectories, a property that can be described as a reduced cellular plasticity.

Analyzing FGF stimulated mESCs showed increased transcriptional noise (Figure 7), potentially allowing cells to explore transcriptional states. I, therefore, asked if the lack of *Med12*-mutant cells populating the intermediate space goes along with lower transcriptional noise. I used VarID2 to quantify the contribution of biological noise to the total noise within the single-cell transcriptomes. *Med12*-mutant cells consistently had lower average biological noise levels in all three cell states, with the strongest difference in Epi cells (Figure 29B). Collectively, these findings suggest that the absence of *Med12* leads to a diminished overall transcriptional output and a reduction in transcriptional variability. These consequences may lead to constrained cellular plasticity during transitions between lineages, fostering an increased segregation between

distinct lineages. This could be potentially caused by the lower exploratory potential of *Med12*-mutant cells due to their slightly less noisy gene expression.

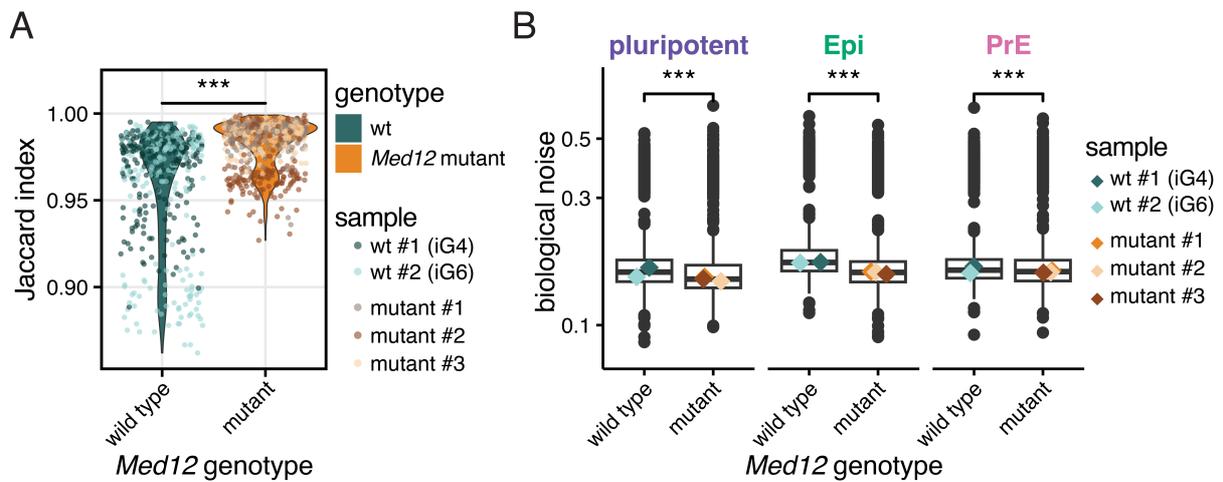


Figure 29: Role of *Med12* in separation between cell types and transcriptional noise.

A Jaccard index to assess cluster stability of wild-type and *Med12*-mutant cells. *** indicates $p < 0.001$, Kolmogorov-Smirnov test.

For clarity, measurements from the two *Med12* wild-type and the three *Med12*-mutant lines were pooled in panels C, D, G, and I, and for the violin plot and statistical test in H.

B Biological noise was determined for each cell state (pluripotency, Epi, and PrE) and genotype (*Med12* wild type and mutant) using VarID2 (Rosales-Alvarez et al., 2023). Medians for single clonal lines are indicated as diamonds. **** indicates $p < 0.0001$, Wilcoxon signed rank test.

5 Discussion

During differentiation cells acquire new functionalities by changing the set of expressed genes. The process of differentiation includes that homogenous cell populations break symmetry to increase the diversity during development. What drives cells to explore different cellular states even in homogenous environments in the early embryo?

In this thesis, transcriptomic analysis of mESCs revealed strong changes in gene expression upon FGF4 stimulation leading to higher noise levels in FGF4-stimulated mESCs. This increased noise was correlated to a decreased transcriptional burst frequency of the FGF target genes. I employed a genetic screen to identify key regulators of developmental gene expression in pluripotent stem cells. The screening approach confirmed that the FGF/ERK and mTOR signaling systems play critical roles in positively and negatively regulating reporter expression, respectively. Additionally, I identified several members of the Mediator and Elongator complexes that are involved in the transcription of *Spry4*. Among these potential transcriptional regulators, *Med12* had the strongest effect on the reporter gene expression.

Focusing on *Med12*, I found that it does not function downstream of a single developmental signal but rather collaborates with multiple signaling systems to regulate gene expression in pluripotent cells and differentiation. Functional assays demonstrated that the loss of *Med12* significantly impaired the re-establishment of a naïve pluripotency gene expression program in colony formation assays. Furthermore, it reduced the propensity to populate transition states in an epiblast and primitive endoderm differentiation paradigm. These phenotypes are consistent with reduced biological noise in *Med12*-mutant cells. Taken together, these results indicate that *Med12* enhances transcriptional changes in pluripotent stem cells. Thereby, *Med12* contributes to cellular plasticity maintenance during differentiation and lineage transitions, which enables cells to efficiently react to changing signaling environments. Together, FGF-signaling and *Med12* cooperate to regulate differentiation potentially by stimulating cells to explore a wider transcriptional space.

5.1 Context of regulators identified in CRISPR screen

5.1.1 Connection between previously described and newly identified regulators

In this thesis, I aimed to explain how FGF target genes are regulated and to investigate transcriptional changes triggered by multiple signaling pathways relevant for development. I presented a CRISPR knockout screen with a readout based on the expression of *Spry4*^{H2B-Venus}. This

reporter is based on the expression of the most upregulated target gene of FGF4 in mESCs *Spry4* and has so far not been used for genetic screening approaches. Together with the flow cytometry-based approach, which detects enriched cells even when *Spry4* levels are mildly affected, this screen is a very sensitive readout for gene perturbations influencing the expression magnitude of *Spry4*. Moreover, *Spry4* generally indicates developmental progression from naïve pluripotency, as it is upregulated in both the primed epiblast and PrE cells in vivo (Morgani et al., 2018). Combined with the robustly designed gRNA library Brie, covering 19,674 protein-coding mouse genes, the screen allowed the identification of multiple functional groups of genes.

The gene list revealed by the screen contains genes both positively and negatively regulators of the FGF/ERK signaling cascade, e.g., receptors or downstream kinases, reflecting the specificity of the screen. Additionally, genes involved in synthesizing extracellular co-factors, e.g., genes involved in the glycosylation and export of heparin sulfates, positively regulated *Spry4* levels. The pooled, genome-wide design of the CRISPR knockout screen causes that the number of neighbor cells with the same genotype is limited by cell division and regular cell reseeded. The occurrence of genes involved in heparin sulfate synthesis as important co-factors for FGF indicates a cell-autonomous function of heparan sulfates. Heparane sulfate synthesis thereby offers a possibility for cells to regulate their responsiveness to external FGF signals. The knockout of non-redundant components of the signaling cascades had the strongest effect on *Spry4* expression, mimicking almost a complete loss of ERK activity in PD03-treated cells.

FGF-signaling was confirmed to be a major regulator of *Spry4* expression. Which proteins link FGF signals with the transcriptional activity of PolII at target gene promoters to implement transcriptional changes? Surprisingly, the CRISPR screen did not reveal any sequence-specific transcriptional regulator downstream of FGF/ERK-signaling. Although FGF signaling genes occurred in a previous screen analyzing perturbations affecting the exit of pluripotency shown by a delayed downregulation of the *Rex1*-GFP reporter, also here, no FGF-dependent transcription factors directly binding to target gene promoter were identified (M. Li et al., 2018). The absence of a potential sequence-specific FGF-dependent transcription factor could result from redundancies. However, the transcription factors ETV4 and ETV5 have previously been proposed to have an FGF-dependent effect on the mESC clonogenicity, even if only one of the factors is mutated (Kalkan et al., 2019). The absence of these genes in the screen, together with the lack of a pre-implantation phenotype of ETV5-mutant mice (Zhang et al., 2021), make these genes unlikely to encode the primary mediator of transcriptional responses to FGF in mESCs. The robustness against loss of ETV5 in the embryo could potentially indicate additional functional redundancy by unknown mechanisms lost in mESCs, for example based on gene expression changes during isolation of mESCs from the ICM. Further cofactors mediating FGF-induced

transcription can be aimed to be identified with locus-specific proteomic approaches at the *Spry4* promoter. Alternative possibilities for how FGF acts on transcription include the possibility of ERK itself transferring to the nucleus and influencing PolII activity directly by phosphorylation (Tee et al., 2014), leaving the mechanism for target gene specificity unanswered. Context-dependent organization of the chromatin and the general transcription machinery might define the promoters influenced by phospho-ERK. To analyze which genes are directly regulated by nuclear ERK binding to PolII, ChIP-Seq with an antibody specific for activated ERK could be used under FGF-stimulation conditions. Comparison with the list of FGF target genes obtained from this thesis can evaluate the generality of direct transcriptional regulation via ppERK during early differentiation.

In contrast to FGF, multiple negative regulators of mTOR signaling components were found as negative regulators of *Spry4* expression. Given the diverse functional involvement in metabolism and proliferation, interference with mTOR signaling might result in altered cell sizes and, therefore, could be the reason for the enrichment during flow cytometry-based sorting. However, knockdowns of *Tsc1/2* and *Flcn* were identified to prevent mESC's commitment in previous screens. *Flcn* directly regulates *Tfe3*, a transcription factor linked to the expression of the pluripotency gene *Esrrb*, and the TSC complex negatively regulates the mTOR complex, which itself inhibits *Tfe3* function (Betschinger et al., 2013; Villegas et al., 2019). The effects of *Tsc1/2* knockdowns were further confirmed to delay the exit of pluripotency in the second screening approach, which also found *Lamtor4* to affect the exit of pluripotency similarly (M. Li et al., 2018). This was connected to mTOR's involvement in β -Catenin activation by GSK3 inhibition. However, I showed, that *Spry4*^{H2B-Venus} expression is independent of Chiron-induced β -Catenin expression, indicating that the mTOR regulators are not enriched in the screen due to disrupted Wnt/ β -Catenin signaling. Since *Spry4* is a target gene of FGF signaling in mESCs, it remains elusive why its expression is upregulated upon perturbations of genes delaying or preventing differentiation and therefore having opposing effects to FGF signaling. One possibility would be additional cross-talk between FGF and mTOR signaling. In human embryonic kidney cells, ERK activity phosphorylates TSC2, resulting in a loss of mTOR inhibition (Ma et al., 2005), possibly mechanistically linking FGF signaling and mTOR activity. A partial dependence of *Spry4*^{H2B-Venus} on FGF/ERK induced mTOR expression potentially explains the finding of negative regulators of mTOR to negatively regulate *Spry4*^{H2B-Venus} expression as well.

Wnt signaling is one of the central signaling systems influencing early differentiation and was also described to influence the expression of *Spry4* (Katoh & Katoh, 2006). However, no hit from the screen indicated a dependence of *Spry4* on the Wnt signaling environment. Additionally, direct stimulation and inhibition of Wnt signaling in mESCs did not affect *Spry4*^{H2B-Venus} expression levels.

This makes *Spry4* unlikely to be a target gene in the used media conditions and cell lines and explains the absence of Wnt target genes in the screen.

The screening process identified several components of FGF and mTOR signaling systems, along with some members of the SWI/SNF and the Mediator complexes. The core Mediator complex is believed to be essential for the expression of most genes in eukaryotic genomes. One of its functions is to alter the nucleosomes around the promoter regions, opening the possibility for cooperativity with the SWI/SNF complex. Mediator subunits were identified to benefit *Pou5f1* expression in naïve pluripotency, thereby contributing to pluripotency maintenance (Kagey et al., 2010). In addition, individual subunits of the Mediator complex have been suggested to regulate gene expression downstream of specific signaling systems. The middle and tail module subunits of the Mediator complex, which were identified to positively regulate *Spry4^{H2B-Venus}* expression, were not yet described to have a state- or signal-specific function. The similar and small effect sizes in the downregulation of the *Spry4^{H2B-Venus}*-reporter in polyclonal knockouts for multiple of these Mediator subunits, excluding *Med12*, could be dependent on an above average dependency of *Spry4* expression on Mediator in general and not be connected to subunit-specific functions. Mechanistically, a strong enhancer regulating *Spry4* could lead to such a strong influence of Mediator.

5.1.2 Limitations of the CRISPR screen approach

The design based on the FGF target gene *Spry4* allowed very sensitive screening for signaling and transcriptional regulators. Using this as a single readout limits the interpretational power. In contrast to other screening approaches, based on pluripotency factors *Pou5f1* or *Zfp42* (Rex1) (Betschinger et al., 2013; Kagey et al., 2010; M. Li et al., 2018), the protein itself does not have a functional role during early cell type transition although *Spry4* is upregulated during these transitions. The functional consequences of gene knockouts influencing *Spry4* have to be further tested, which is limited to *Med12* within this thesis. The high sensitivity leads to the enrichment of reasonable functional groups, as discussed above. However, hits among ribosomal subunits and with mitochondrial function imply the possibility of significant hits within genes important for homeostatic functions and proliferation, rather than signaling-related transcriptional regulation during cell type transitions. Furthermore, this screen design does not include a condition of active stimulation leading to cell type transitions. Genes that only regulate gene expression in response to differentiation cues cannot be expected to be detected. This could be addressed with screens focusing on readouts for early differentiation. The issue of redundancy could be addressed by further screening with a combinatorial approach, potentially with single cell resolution (Replogle et al., 2020) focusing on transcriptional regulators.

5.2 Functional characterization of *Med12*

Med12 was identified in the CRISPR screen as the transcriptional regulator influencing *Spry4^{H2B-Venus}*-expression the strongest. In previous studies *Med12* was shown to be essential for the development of the peri-implantation embryo, where its absence disrupted Wnt-signaling (Rocha et al., 2010). Moreover, *Med12* was directly linked to pluripotency maintenance via interaction with promoters of pluripotency factors (Apostolou et al., 2013; Tutter et al., 2009) and is part of the kinase module of the Mediator Complex, which can be inhibited to maintain pluripotency in absence of 2i (Lynch et al., 2020). I therefore proceed with a more extensive analysis of the role of *Med12* in pluripotency and during signaling-induced differentiation including the exit of pluripotency and PrE differentiation.

5.2.1 Role of *Med12* in pluripotency

Med12 influences gene expression in pluripotency states and transitions. In mESCs, a direct relationship between *Med12* and pluripotency was proposed based on the finding that *Med12* binds NANOG, which is part of a group of transcription factors characteristic of pluripotency (Tutter et al., 2009). A knockdown of *Nanog* led to similar gene expression changes as knockdown of *Med12* (Tutter et al., 2009). Furthermore, *Med12* bound *Nanog*-dependent gene promoters in pluripotency and was absent during differentiation, following the dynamics of NANOG (Apostolou et al., 2013; Tutter et al., 2009). A contradicting study found that *Nanog* activity in gene regulation does not depend on the presence of *Med12* (Rocha et al., 2010). The data presented in this thesis suggest only a small effect of the presence of *Med12* on *Nanog* since in pluripotency media, proliferation and pluripotency characteristics were maintained, and naïve pluripotency-dependent genes were only mildly downregulated compared to previous experiments (Tutter et al., 2009). However, these experiments were performed in ES + LIF media compared to the differentiation in molecularly defined N2B27 used here. Moreover, the acute downregulation with siRNAs might contribute to stronger effects since compensation mechanisms could be activated in long-term culture of *Med12*-mutant cells.

The expression of *Nanog* as one of the key pluripotency factors was shown to be dependent on the activity of an enhancer of unusually large size and Mediator occupation. These super-enhancers were linked to crucial identity genes in different cell states (Whyte et al., 2013; Zamudio et al., 2019). The dependence of cell state-specific genes on these enhancers and the activity of Mediator to bridge these enhancers to the core promoter could be a possible explanation for the effects of *Med12* mutation on pluripotency maintenance. *Med12*'s function to bind non-coding RNAs at enhancers, in order to tether the mediator complex associated to

expressed genes and enhancers (Lai et al., 2013), could explain why *Med12* has a stronger effect than other Mediator subunits. However, the effect magnitude of the loss of *Med12* was different between genes. This is unlikely to reflect solely the genes' dependence on super-enhancers since the expression of the inducible transgene *GATA6* is strongly affected by the loss of *Med12*, while the transgenic introduction at random positions in the genome make corresponding super-enhancer activity unlikely. Therefore, *Med12* activity is not only dependent on the presence of super-enhancers.

5.2.2 *Med12* cooperates with signaling during pluripotency transitions

As described above, previous studies suggest signal-specific functions of the Mediator subunit *Med12*. In this thesis, *Med12* promoted expression of *Spry4* in ES + LIF, as shown by the screen and polyclonal knockout lines. The monoclonal *Med12*-mutant lines expressed lower naïve pluripotency marker levels and downregulated these genes more slowly upon differentiation signals. To test if these effects depend on specific pluripotency-associated signals, I employed a footprinting analysis comparing effects on defined sets of target genes for mTOR, Notch, FGF and Wnt-signaling. However, this did not predict any reliable dependency on a single signaling system. The high score for similarity with mTOR signaling could not be confirmed through any epistatic relation to *Med12*, tested by double knockout experiments. This discrepancy potentially indicates the high connectivity of mTOR components to other signals and thereby prevents the definition of mTOR-specific target genes. Previously, *Med12* was linked to Wnt/ β -Catenin and Wnt/PCP signaling activity in early mammalian development. Specifically, *Med12*-mutant embryos failed to undergo early gastrulation and did not establish an anterior-posterior axis, since they lack an AVE population (Rocha et al., 2010). However, *Med12*-mutant mESCs from this study established an AVE identity in a 2D system. Notably, *Med12*-mutant embryos appeared different from wild-type embryos in size and structure, opening the possibility that embryonic development is affected even earlier and disrupted Wnt signaling is potentially a consequence of these effects. Together, these findings suggest a strong context-dependency regarding the functional connection between *Med12* and Wnt.

Med12 is an essential structural component of the Mediator kinase module, which was recently linked to FGF/ERK signaling and maintenance of pluripotency in mouse and human stem cells. Similar to 2i treatment, cells were kept in a pluripotent state via inhibition of the kinase function of the Mediator kinase module. The inhibitor used by Lynch et al. (2020) was described to be specific for CDK8 and 19 and replacing CKD8/19 with a kinase dead mutant had similar effects. CDK8/19 inhibited cells were similar to cells in 2i with respect to their ability for colony formation, gene expression, and contribution to mouse embryos. MEK inhibition resulted in

CDK8/19 inhibition (Lynch et al., 2020). In this thesis, however, FGF target genes in *Med12* mutant cells only showed a median 6 to 7% reduction in their foldchange compared to wild-type cells. Moreover, *Med12*-mutants express lower naïve pluripotency marker levels than wild-type cells in the presence of 2i. This could be explained by the loss of *Med12* preventing the formation of the Mediator kinase module. A complete loss of the formation of the kinase module could have different effects than its inhibition since this maintains its structural presence. This is supported by Lynch et al. (2019), finding that the presence of the inactivated CDK8/19 is required for the maintenance of pluripotency. Furthermore, there could be kinase module-independent functions of MED12. A cytoplasmic function directly at the intracellular domain of TGFβ-receptors was proposed for *Med12* in human lung cancer (Huang et al., 2012). However, in the *Med12*-mutant mESCs, ppERK levels were not systematically affected. This indicates a function of *Med12* during transcriptional regulation and not upstream of the phosphorylation of ERK. In hemopoietic stem cells, the deletion of MED12 destabilizes the chromatin signature at lineage-specific enhancers, resulting in a loss of their stem cell potency (Aranda-Orgilles et al., 2016). Similar roles of *Med12* in chromatin modifications during pluripotency could explain the cooperativity of *Med12* with different signaling systems in pluripotency. Both possibilities, a kinase module independent function as well as the requirement of the presence of the inactive kinase module to maintain pluripotency, are compatible with the absence of other kinase module components in the *Spry4^{H2B-Venus}*-based CRISPR screen. To distinguish between kinase module dependent and independent functions of *Med12*, the effect of a Mediator kinase module inhibitor on *Med12*-mutant cells should be tested. The clonogenicity assay as a sensitive readout for lineage commitment could be used to compare the effect of CDK8/19 inhibition on *Med12*-mutant and wild-type cells.

Generally, the loss of *Med12* is well tolerated by mESCs in stable conditions. Cell type transitions are accomplished by *Med12*-mutant cells, however, sometimes with different dynamics and plasticity. These mild phenotypes could indicate that *Med12*'s functions are only partly lost. While incomplete knockdown was excluded by western blot analysis, redundancy could play a role. Its paralog *Med12L* is upregulated 2 to 3-fold on the mRNA level upon loss of *Med12*. However, during differentiation or stimulation, it is slightly downregulated, arguing against a function during cell type transitions. Mutation of *Med12L* in a *Med12*-mutant background further lowered *Spry4^{H2B-Venus}* expression in ES + LIF conditions, indicating some redundancy in the regulation of single genes. Mechanistically, it remains speculative which regulating functions of *Med12*, *Med12L* is able to overtake, but 67% identity with *Med12* including important domains indicate a high potential (Luyties & Taatjes, 2022). Additionally, for the long-term disruption of other Mediator subunits, a feedback loop via CDK9 was recently described to compensate the PolII activation by Mediator

(Jaeger et al., 2020). The described degradation tag constructs could be used to investigate similar compensation mechanisms for *Med12*.

5.2.3 mESC populations buffer PrE differentiation against loss of *Med12*

The sensitivity of individual gene expression to the loss of *Med12* is variable, as evidenced by the increased sensitivity of the *Sprouty4* locus compared to the most of FGF target genes. On the level of the expression of single gene, the dependency on *Med12* extends to exogenous transgenes like the inducible GATA6-mCherry construct used in primitive endoderm differentiation. Specifically, the loss of MED12 altered the dynamics of GATA6 expression during doxycycline induction. Shortening the induction time in the wild type aligned the total GATA6 abundance levels integrated over time and the GATA6 levels at the end of the induction period. Still, it could not equalize the maximum average expression levels. This effect on gene expression dynamics reflects the phenotype of *Med12*-mutant cells to regulate gene expression less efficiently in response to stimuli during the exit of pluripotency. However, this is surprising since the expression of the induced transgene depends on the doxycycline-dependent binding of a modified version of the tetracycline (Tet) repressor protein to the promoter region of the GATA6 transgene. The introduced binding site is directly upstream of a minimal CMV promoter (Schröter et al., 2015); therefore, it is not comparable to an enhancer. Moreover, the Tet repressor protein is derived from *Escherichia Coli* bacteria (Baron & Bujard, 2000), which do not have a conserved mediator complex.

In wild-type cells, the ratio between Epi and PrE cells differentiated from mESCs after doxycycline induction is robust against variations in GATA induction duration and levels within a specific range (Raina et al., 2021). In *Med12*-mutant cells decreased GATA induction levels led to reduced proportions of PrE cells. This indicates that the minimal threshold of GATA induction is not reached in some *Med12*-mutant cells, which, therefore, stay refractory to PrE differentiation, even when extended doxycycline induction and supplemented with FGF4. This threshold seems to be set by the maximal expression strength and not by the integrated total expression of the induced transgene.

As discussed above, FGF signaling was still functional in *Med12*-mutant cells during the exit of pluripotency. Confirming this, during PrE differentiation FGF supplementation rescued the potential to fully convert to PrE cells of cells with high GATA6 induction levels, an effect similarly observed in wild-type cells for aligned GATA6 induction levels. Despite quantitative defects in the regulation of individual genes, no significant defects are observed in the acquisition of early differentiated fates. Marker gene expression levels and transcriptomic analysis reveal that the

composition of cellular transcriptomes remains relatively consistent between *Med12*-mutant and wild-type cells, indicating that intracellular regulatory networks can buffer against variable transcription efficiencies. The remaining differences between *Med12*-mutant and wild-type cells were mainly cell state independent, indicating *Med12* impacts the transition and commitment on trajectories between states rather than influencing gene expression during stable states.

5.2.4 *Med12* contributes to plasticity during differentiation

In this thesis, the investigation of the functions of *Med12* during lineage transitions revealed that *Med12*-mutant cells exhibit reduced transcriptional plasticity. This reduction is evident in slower downregulation of pluripotency genes, decreased capacity to revert to naïve pluripotency in colony-forming assays, and a diminished tendency to populate transition states between an epiblast and a primitive endoderm identity. An explanation for the higher proportion of cells between the defined cell types in the wild-type could be cells switching between Epi and PrE precursor cell states late during the trajectory, which manifests a prime example of plastic behavior during differentiation. Following endogenous GATA levels, which indicate the differentiation trajectory of a single cell early, with a GATA-reporter line (Raina et al., 2021) could further prove the requirement of *Med12* for these trajectory switches. The role of *Med12* to maintain plasticity and responsiveness to changing signaling environments might be related to the 3D genome organization around *Med12*-targeted promoters. In a screen focused on transcription and chromatin regulators during maintenance of pluripotency various Mediator subunits, including *Med12*, were identified (Kagey et al., 2010). The study proposed that interactions between Mediator and cohesin contribute to genome folding and efficient enhancer-promoter interactions. I speculate that the observed cellular phenotypes stem from the impaired ability of *Med12*-mutant cells to reconfigure chromatin in response to changing signaling environments. This could be tested by tracking the genome organization with Hi-C-based methods during differentiation. Further experiments with *Med12*-mutant cells could test how the decreased plasticity influences the developmental potential to form a multitude of cell types. Specifically, to analyze the generality of the effect of *Med12*, comparison of wild-type and *Med12*-mutant gastruloids could reveal if the reduced plasticity generally results in clearer cell type separation during differentiation of a multitude of cell types. In comparison to the investigated Epi/PrE-differentiation paradigm, gastruloids generation would be independent of the induction of a single transgene. Finally, it is unclear if the potency of *Med12* to contribute to cellular plasticity is actively regulated during embryonic development. CDK8 protein levels are higher in PrE cells than in the early Epi cells in vivo at embryonic day 4.5, while this reverts in the later Epi at embryonic day 5.5 (Lynch et al., 2020). If *Med12* acts independent of the kinase module,

differential expression could be an attractive solution to influence plasticity, correlating to the high plasticity of PrE cells (Grabarek et al., 2012).

5.3 FGF and *Med12* regulate noise during differentiation

Differentiation signals, in conjunction with regulatory DNA sequences and transcription factors, shape gene-specific transcription dynamics. The transcriptional bursting dynamics of target genes can be described with the transcriptional burst size and frequency, which together control the transcriptional noise levels. Here, I used droplet-based sequencing with high sequencing depth to estimate transcriptional burst parameters genome-wide. Previously, this was only performed with array-based single-cell sequencing techniques developed to analyze large fractions of the total mRNA content of a cell (Larsson et al., 2019; Ochiai et al., 2020). However, comparisons of the obtained mRNA count distribution between droplet-based single-cell sequencing and in situ HCR showed high mRNA counting efficiency of droplet-based RNA sequencing. For genes expressed above a certain threshold level, the obtained mRNA count distributions enabled efficient burst parameter estimation. I showed FGF target genes are characterized by lower burst frequencies and therefore higher noise levels, since lower burst frequencies cause higher noise levels if the average expression is kept constant by the burst frequency (Figure 5). This is coherent with a previous report, in which MEK inhibition reduced the transcriptional noise globally (Ochiai et al., 2020). This increased transcriptional noise potentially enables the cells to explore differentiation trajectories. A more direct link between a cells' ability to explore cell trajectories and to differentiate by increased transcriptional noise comes from the analysis of a single gene: Stochastic expression of *Nanog* and high variability in NANOG protein levels was previously linked to initiation of differentiation (Abranches et al., 2014; Ochiai et al., 2014).

The noise levels of cells and therefore their explorative potential has to be tightly regulated. In this thesis, *Med12* was identified as a candidate for this. The FGF target gene *Spry4*, which is transcribed with high noise levels, is less upregulated in *Med12*-mutant cells. Potentially, this could be caused by lower burst sizes of *Spry4* expression in the absence of *Med12*. Moreover, *Med12*-mutant cells show generally decreased biological noise levels during pluripotency and differentiation. Supporting this notion, *Med12*'s localization correlates with the burst size of bound genes (Ochiai et al., 2020), contributing to higher gene expression noise in wild-type cells. Mechanistically, *Med12*'s functions in activating transcriptional pause release or elongation, probably via the proposed activation of the super elongation complex (Galbraith et al., 2013), offers a possibility how *Med12* could induce larger transcriptional burst sizes at targeted genes.

Together, both *Med12* and FGF promote differentiation through the induction of higher noise levels and thereby enlarging the exploratory potential of cells. This common functional role provides an explanation for the specific enrichment of FGF signaling components and *Med12* in the CRISPR screen for transcriptional regulation of *Spry4*. Yet, it is unclear, how general this functional cooperation of *Med12* and FGF is to promote noisy gene expression at developmental promoters. Measuring transcriptional noise levels and gene expression with single cell resolution in response to FGF signaling, directly comparing *Med12*-mutant with wild-type cells would help to resolve this question.

5.4 Conclusions and future perspectives

The analysis of the effects of FGF signaling with single-cell resolution showed that FGF is one factor that drives cells to explore different transcriptional states by contributing to higher transcriptional noise levels in mESCs. The highly stochastic expression of FGF target genes, which have a lower transcriptional burst frequency, causes heterogeneity essential for cellular symmetry breaking, for example, in the inner cell mass. Surprisingly, no sequence-specific transcription factor was identified as a molecular mediator of these FGF-induced transcriptional changes, even in a highly specific, genome-wide screen. This could imply the possibility that the function of FGF to induce transcriptional noise during the exit of pluripotency is independent of any sequence-specific transcription factor. Rather, FGF target genes could be activated by regulating transcriptional activity parameters via affecting general transcription factors at a context-dependent set of genes. One factor of the general transcriptional machinery is *Med12*, which was identified to influence the FGF target gene *Spry4* with the genome-wide screen. While the average gene expression changes upon FGF signaling in multiple differentiation paradigms did not change majorly in *Med12*-mutant cells compared to the wild-type, *Med12* mutant cells showed lower biological noise levels in both pluripotency and upon differentiation to PrE- and Epi-cells. This positions *Med12* as an important factor for differentiation in parallel to signaling events. Functionally, higher noise levels in *Med12* wild-type cells contribute to the cell's ability to efficiently react to changing environmental signals resulting in plasticity during differentiation. In conclusion, FGF-signaling and *Med12* mark two important factors to enable differentiation from pluripotent cells. Recruiting *Med12* to promoters could be a general mechanism to increase the expression noise of the signaling-targeted genes by potentially increasing the transcriptional burst size while keeping the burst frequency low.

For both FGF and *Med12*, it remained unclear what defines their target genes and how much overlap there is between them in the context of early differentiation. The low impact on the

average expression of FGF target genes does not exclude that the FGF-induced effect of increased transcriptional noise is implemented by *Med12*, especially considering the plethora of proposed mechanisms of action for *Med12*, which could lead to higher transcriptional noise, while not affecting the average expression change. To understand this potential cooperation better, transcriptional noise levels have to be compared between *Med12*-mutant and wild-type cells globally or at beforehand identified promoters affected by both factors. Therefore, the performed single-cell RNA sequencing experiment with high capturing efficiency for FGF stimulation could be extended to include *Med12*-mutant cells and combined with a ChIP-seq experiment to identify the promoters MED12 binds to.

Collectively, these experiments will further test the mechanisms of cooperativity between FGF-signaling and *Med12* as an example of how the effects of signaling are influenced by the general transcriptional machinery. Extending this idea to multiple signaling systems and general transcriptional regulators can further the understanding of how the complexity of an embryo and the transcriptional space of diverse cell types can be established robustly and within a short time frame.

6 References

- Abranches, E., Guedes, A. M. V., Moravec, M., Maamar, H., Svoboda, P., Raj, A., & Henrique, D. (2014). Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. *Development*, *141*(14), 2770–2779. <https://doi.org/10.1242/dev.108910>
- Akagi, T., Kuure, S., Uranishi, K., Koide, H., Costantini, F., & Yokota, T. (2015). ETS-related transcription factors Etv4 and Etv5 are involved in proliferation and induction of differentiation-associated genes in embryonic stem (ES) cells. *Journal of Biological Chemistry*, *290*(37), 22460–22473. <https://doi.org/10.1074/jbc.M115.675595>
- Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L., Hill, D. P., Lee, R., Mi, H., Moxon, S., Mungall, C. J., Muruganugan, A., Mushayahama, T., Sternberg, P. W., Thomas, P. D., Van Auken, K., ... Westerfield, M. (2023). The Gene Ontology knowledgebase in 2023. *Genetics*, *224*(1). <https://doi.org/10.1093/GENETICS/IYAD031>
- Amadei, G., Handford, C. E., Qiu, C., De Jonghe, J., Greenfeld, H., Tran, M., Martin, B. K., Chen, D. Y., Aguilera-Castrejon, A., Hanna, J. H., Elowitz, M. B., Hollfelder, F., Shendure, J., Glover, D. M., & Zernicka-Goetz, M. (2022). Embryo model completes gastrulation to neurulation and organogenesis. *Nature*, *610*(7930), 143–153. <https://doi.org/10.1038/s41586-022-05246-3>
- Apostolou, E., Ferrari, F., Walsh, R. M., Bar-Nur, O., Stadtfeld, M., Cheloufi, S., Stuart, H. T., Polo, J. M., Ohsumi, T. K., Borowsky, M. L., Kharchenko, P. V., Park, P. J., & Hochedlinger, K. (2013). Genome-wide chromatin interactions of the nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell*, *12*(6), 699–712. <https://doi.org/10.1016/j.stem.2013.04.013>
- Aranda-Orgilles, B., Saldaña-Meyer, R., Wang, E., Trompouki, E., Fassl, A., Lau, S., Mullenders, J., Rocha, P. P., Raviram, R., Guillamot, M., Sánchez-Díaz, M., Wang, K., Kayembe, C., Zhang, N., Amoasii, L., Choudhuri, A., Skok, J. A., Schober, M., Reinberg, D., ... Aifantis, I. (2016). MED12 Regulates HSC-Specific Enhancers Independently of Mediator Kinase Activity to Control Hematopoiesis. *Cell Stem Cell*, *19*(6), 784–799. <https://doi.org/10.1016/j.stem.2016.08.004>
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. <https://doi.org/10.1038/75556>
- Athanasouli, P., Balli, M., De Jaime-Soguero, A., Boel, A., Papanikolaou, S., van der Veer, B. K., Janiszewski, A., Vanhessche, T., Francis, A., El Laithy, Y., Nigro, A. Lo, Aulicino, F., Koh, K. P., Pasque, V., Cosma, M. P., Verfaillie, C., Zwijsen, A., Heindryckx, B., Nikolaou, C., & Lluís, F. (2023). The Wnt/TCF7L1 transcriptional repressor axis drives primitive endoderm formation by antagonizing naive and formative pluripotency. *Nature Communications*, *14*(1), 1–19. <https://doi.org/10.1038/s41467-023-36914-1>
- Baillie-Benson, P., Moris, N., & Martinez Arias, A. (2020). Pluripotent stem cell models of early mammalian development. *Current Opinion in Cell Biology*, *66*, 89–96. <https://doi.org/10.1016/j.CEB.2020.05.010>
- Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C. J., Mclachlan, H., Klevernic, I., Arthur, J. S. C., Alessi, D. R., & Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. *The Biochemical Journal*, *408*(3), 297–315. <https://doi.org/10.1042/BJ20070797>
- Balamotis, M. A., Pennella, M. A., Stevens, J. L., Wasylyk, B., Belmont, A. S., & Berk, A. J. (2009). Complexity in transcription control at the activation domain-Mediator interface. *Science Signaling*, *2*(69). <https://doi.org/10.1126/scisignal.1164302>

- Baron, U., & Bujard, H. (2000). Tet repressor-based system for regulated gene expression in eukaryotic cells: Principles and advances. *Methods in Enzymology*, 327, 401–421. [https://doi.org/10.1016/S0076-6879\(00\)27292-3](https://doi.org/10.1016/S0076-6879(00)27292-3)
- Barrett, S. D., Bridges, A. J., Dudley, D. T., Saltiel, A. R., Fergus, J. H., Flamme, C. M., Delaney, A. M., Kaufman, M., LePage, S., Leopold, W. R., Przybranowski, S. A., Sebolt-Leopold, J., Van Becelaere, K., Doherty, A. M., Kennedy, R. M., Marston, D., Howard, W. A., Smith, Y., Warmus, J. S., & Tecle, H. (2008). The discovery of the benzhydroxamate MEK inhibitors CI-1040 and PD 0325901. *Bioorganic & Medicinal Chemistry Letters*, 18(24), 6501–6504. <https://doi.org/10.1016/J.BMCL.2008.10.054>
- Bartman, C. R., Hsu, S. C., Hsiung, C. C. S., Raj, A., & Blobel, G. A. (2016). Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell*, 62(2), 237–247. <https://doi.org/10.1016/J.MOLCEL.2016.03.007>
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., & Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Molecular Cell*, 2(4), 437–445. [https://doi.org/10.1016/S1097-2765\(00\)80143-4](https://doi.org/10.1016/S1097-2765(00)80143-4)
- Bessonard, S., Mot, L. De, Gonze, D., Barriol, M., Dennis, C., Goldbeter, A., Dupont, G., & Chazaud, C. (2014). Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Development*, 141(19), 3637–3648. <https://doi.org/10.1242/DEV.109678>
- Betschinger, J., Nichols, J., Dietmann, S., Corrin, P. D., Paddison, P. J., & Smith, A. (2013). Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell*, 153(2), 335–347. <https://doi.org/10.1016/j.cell.2013.03.012>
- Bulut-Karslioglu, A., Biechele, S., Jin, H., MacRae, T. A., Hejna, M., Gertsenstein, M., Song, J. S., & Ramalho-Santos, M. (2016). Inhibition of mTOR induces a paused pluripotent state. *Nature*, 540(7631), 119–123. <https://doi.org/10.1038/nature20578>
- Carlini, V., Gretarsson, K. H., & Hackett, J. A. (2021). Genome-scale CRISPR screening for regulators of cell fate transitions. In *Methods in Molecular Biology* (Vol. 2214, pp. 91–108). Humana Press Inc. https://doi.org/10.1007/978-1-0716-0958-3_7
- Carter, M. G., Sharov, A. A., VanBuren, V., Dudekula, D. B., Carmack, C. E., Nelson, C., & Ko, M. S. (2005). Transcript copy number estimation using a mouse whole-genome oligonucleotide microarray. *Genome Biology*, 6(7), R61. <https://doi.org/10.1186/GB-2005-6-7-R61>
- Chazaud, C., & Rossant, J. (2006). Disruption of early proximodistal patterning and AVE formation in *Apc* mutants. *Development*, 133(17), 3379–3387. <https://doi.org/10.1242/DEV.02523>
- Chazaud, C., & Yamanaka, Y. (2016). Lineage specification in the mouse preimplantation embryo. *Development*, 143(7), 1063–1074. <https://doi.org/10.1242/DEV.128314>
- Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., & Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development*, 145(12), 1–10. <https://doi.org/10.1242/DEV.165753>
- Clevers, H., & Nusse, R. (2012). Wnt/ β -Catenin Signaling and Disease. *Cell*, 149(6), 1192–1205. <https://doi.org/10.1016/J.CELL.2012.05.012>
- Cline, G. W., Johnson, K., Regittnig, W., Perret, P., Tozzo, E., Xiao, L., Damico, C., & Shulman, G. I. (2002). Effects of a Novel Glycogen Synthase Kinase-3 Inhibitor on Insulin-Stimulated Glucose Metabolism in Zucker Diabetic Fatty (fa/fa) Rats. *Diabetes*, 51(10), 2903–2910. <https://doi.org/10.2337/DIABETES.51.10.2903>
- Corden, J. L. (2013). RNA polymerase II C-terminal domain: Tethering transcription to transcript and template. *Chemical Reviews*, 113(11), 8423–8455. <https://doi.org/10.1021/cr400158h>
- Daniels, D. L., Ford, M., Schwinn, M. K., Benink, H., D.Galbraith, M., Amunugama, R., Jones, R., Allen, D., Okazaki, N., Yamakawa, H., Miki, F., Nagase, T., M.Espinosa, J., & Urh, M. (2013). Mutual Exclusivity of MED12/MED12L, MED13/13L, and CDK8/19 Paralogs Revealed within the

- CDK-Mediator Kinase Module. *Journal of Proteomics & Bioinformatics*, 0, 1–7. <https://doi.org/10.4172/JPB.S2-004>
- Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., & Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, 34(2), 184–191. <https://doi.org/10.1038/nbt.3437>
- Donner, A. J., Ebmeier, C. C., Taatjes, D. J., & Espinosa, J. M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nature Structural and Molecular Biology*, 17(2), 194–201. <https://doi.org/10.1038/nsmb.1752>
- Edelstein, A., Amodaj, N., Hoover, K., Vale, R., & Stuurman, N. (2010). Computer control of microscopes using μ Manager. *Current Protocols in Molecular Biology*, Chapter 14. <https://doi.org/10.1002/0471142727.MB1420S92>
- Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297(5584), 1183–1186. <https://doi.org/10.1126/science.1070919>
- Ershov, D., Phan, M. S., Pylvänäinen, J. W., Rigaud, S. U., Le Blanc, L., Charles-Orszag, A., Conway, J. R. W., Laine, R. F., Roy, N. H., Bonazzi, D., Duménil, G., Jacquemet, G., & Tinevez, J. Y. (2022). TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nature Methods*, 19(7), 829–832. <https://doi.org/10.1038/s41592-022-01507-1>
- Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292(5819), 154–156. <https://doi.org/10.1038/292154a0>
- Eychenne, T., Novikova, E., Barrault, M. B., Alibert, O., Boschiero, C., Peixeiro, N., Cornu, D., Redeker, V., Kuras, L., Nicolas, P., Werner, M., & Soutourina, J. (2016). Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture. *Genes & Development*, 30(18), 2119–2132. <https://doi.org/10.1101/GAD.285775.116>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/J.PATREC.2005.10.010>
- Frankenberg, S., Gerbe, F., Bessonard, S., Belville, C., Pouchin, P., Bardot, O., & Chazaud, C. (2011). Primitive Endoderm Differentiates via a Three-Step Mechanism Involving Nanog and RTK Signaling. *Developmental Cell*, 21(6), 1005–1013. <https://doi.org/10.1016/J.DEVCEL.2011.10.019>
- Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J. I., & Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes & Development*, 16(7), 784–789. <https://doi.org/10.1101/GAD.968802>
- Fukaya, T., Lim, B., & Levine, M. (2016). Enhancer Control of Transcriptional Bursting. *Cell*, 166(2), 358–368. <https://doi.org/10.1016/j.cell.2016.05.025>
- Galbraith, M. D., Allen, M. A., Bensard, C. L., Wang, X., Schwinn, M. K., Qin, B., Long, H. W., Daniels, D. L., Hahn, W. C., Dowell, R. D., & Espinosa, J. M. (2013). XHIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. *Cell*, 153(6), 1327. <https://doi.org/10.1016/j.cell.2013.04.048>
- Gangloff, Y.-G., Mueller, M., Dann, S. G., Svoboda, P., Sticker, M., Spetz, J.-F., Um, S. H., Brown, E. J., Cereghini, S., Thomas, G., & Kozma, S. C. (2004). Disruption of the Mouse mTOR Gene Leads to Early Postimplantation Lethality and Prohibits Embryonic Stem Cell Development. *Molecular and Cellular Biology*, 24(21), 9508–9516. <https://doi.org/10.1128/MCB.24.21.9508-9516.2004>
- Ge, S. X., Jung, D., Jung, D., & Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, 36(8), 2628–2629. <https://doi.org/10.1093/BIOINFORMATICS/BTZ931>
- Göke, J., Chan, Y. S., Yan, J., Vingron, M., & Ng, H. H. (2013). Genome-wide Kinase-Chromatin Interactions Reveal the Regulatory Network of ERK Signaling in Human Embryonic Stem Cells. *Molecular Cell*, 50(6), 844–855. <https://doi.org/10.1016/j.molcel.2013.04.030>

- Grabarek, J. B., Zyzyńska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J., Hadjantonakis, A. K., & Plusa, B. (2012). Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. *Development*, *139*(1), 129–139. <https://doi.org/10.1242/DEV.067702>
- Grünberg, S., Henikoff, S., Hahn, S., & Zentner, G. E. (2016). Mediator binding to UAS s is broadly uncoupled from transcription and cooperative with TFIID recruitment to promoters. *The EMBO Journal*, *35*(22), 2435–2446. <https://doi.org/10.15252/embj.201695020>
- Guo, G., Huss, M., Tong, G. Q., Wang, C., Li Sun, L., Clarke, N. D., & Robson, P. (2010). Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Developmental Cell*, *18*(4), 675–685. <https://doi.org/10.1016/j.DEVCEL.2010.02.012>
- Gurdon, J. B. (1962). The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles. *Development*, *10*(4), 622–640. <https://doi.org/10.1242/DEV.10.4.622>
- Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., & Satija, R. (2023). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology*, 1–12. <https://doi.org/10.1038/s41587-023-01767-y>
- Hernand, D. (2020). Anticodon wobble uridine modification by elongator at the crossroad of cell signaling, differentiation, and diseases. In *Epigenomes* (Vol. 4, Issue 2). MDPI. <https://doi.org/10.3390/epigenomes4020007>
- Hooper, M., Hardy, K., Handyside, A., Hunter, S., & Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature*, *326*(6110), 292–295. <https://doi.org/10.1038/326292A0>
- Hoppe, C., Bowles, J. R., Minchington, T. G., Sutcliffe, C., Upadhyai, P., Rattray, M., & Ashe, H. L. (2020). Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Developmental Cell*, *54*(6), 727–741. <https://doi.org/10.1016/j.DEVCEL.2020.07.007>
- Hoshino, H., Shioi, G., & Aizawa, S. (2015). AVE protein expression and visceral endoderm cell behavior during anterior–posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. *Developmental Biology*, *402*(2), 175–191. <https://doi.org/10.1016/j.YDBIO.2015.03.023>
- Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S. H., & Elledge, S. J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes & Development*, *23*(7), 837–848. <https://doi.org/10.1101/GAD.1769609>
- Huang, S., Hölzel, M., Knijnenburg, T., Schlicker, A., Roepman, P., McDermott, U., Garnett, M., Grenrum, W., Sun, C., Prahallad, A., Groenendijk, F. H., Mittempergher, L., Nijkamp, W., Neefjes, J., Salazar, R., Ten Dijke, P., Uramoto, H., Tanaka, F., Beijersbergen, R. L., ... Bernards, R. (2012). MED12 controls the response to multiple cancer drugs through regulation of TGF- β receptor signaling. *Cell*, *151*(5), 937–950. <https://doi.org/10.1016/j.cell.2012.10.035>
- Inoki, K., Ouyang, H., Zhu, T., Lindvall, C., Wang, Y., Zhang, X., Yang, Q., Bennett, C., Harada, Y., Stankunas, K., Wang, C. yu, He, X., MacDougald, O. A., You, M., Williams, B. O., & Guan, K. L. (2006). TSC2 Integrates Wnt and Energy Signals via a Coordinated Phosphorylation by AMPK and GSK3 to Regulate Cell Growth. *Cell*, *126*(5), 955–968. <https://doi.org/10.1016/j.CELL.2006.06.055>
- Jaeger, M. G., Schwalb, B., Mackowiak, S. D., Velychko, T., Hanzl, A., Imrichova, H., Brand, M., Agerer, B., Chorn, S., Nabet, B., Ferguson, F. M., Müller, A. C., Bergthaler, A., Gray, N. S., Bradner, J. E., Bock, C., Hnisz, D., Cramer, P., & Winter, G. E. (2020). Selective Mediator dependence of cell-type-specifying transcription. *Nature Genetics* *2020* *52*:7, *52*(7), 719–727. <https://doi.org/10.1038/s41588-020-0635-0>
- Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., Van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J., & Young, R. A. (2010). Mediator

- and cohesin connect gene expression and chromatin architecture. *Nature*, 467(7314), 430–435. <https://doi.org/10.1038/nature09380>
- Kalkan, T., Bornelöv, S., Mulas, C., Diamanti, E., Lohoff, T., Ralser, M., Middelkamp, S., Lombard, P., Nichols, J., & Smith, A. (2019). Complementary Activity of ETV5, RBPJ, and TCF3 Drives Formative Transition from Naive Pluripotency. *Cell Stem Cell*, 24(5), 785–801. <https://doi.org/10.1016/j.stem.2019.03.017>
- Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H. J., Nett, I., Marks, H., Walker, R., Stunnenberg, H. G., Lilley, K. S., Nichols, J., Reik, W., Bertone, P., & Smith, A. (2017). Tracking the embryonic stem cell transition from ground state pluripotency. *Development*, 144(7), 1221–1234. <https://doi.org/10.1242/dev.142711>
- Kang, M., Garg, V., & Hadjantonakis, A. K. (2017). Lineage Establishment and Progression within the Inner Cell Mass of the Mouse Blastocyst Requires FGFR1 and FGFR2. *Developmental Cell*, 41(5), 496–510. <https://doi.org/10.1016/j.devcel.2017.05.003>
- Kang, M., Piliszek, A., Artus, J., & Hadjantonakis, A. K. (2013). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Development*, 140(2), 267–279. <https://doi.org/10.1242/DEV.084996>
- Katoh, Y., & Katoh, M. (2006). FGF signaling inhibitor, SPRY4, is evolutionarily conserved target of WNT signaling pathway in progenitor cells. *International Journal of Molecular Medicine*, 17(3), 529–532. <https://doi.org/10.3892/ijmm.17.3.529>
- Kelly, S. J. (1977). Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *The Journal of Experimental Zoology*, 200(3), 365–376. <https://doi.org/10.1002/JEZ.1402000307>
- Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, 37(8), 907–915. <https://doi.org/10.1038/S41587-019-0201-4>
- Kim, J. K., & Marioni, J. C. (2013). Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. *Genome Biology*, 14(1), 1–12. <https://doi.org/10.1186/gb-2013-14-1-r7>
- Knuesel, M. T., Meyer, K. D., Donner, A. J., Espinosa, J. M., & Taatjes, D. J. (2009). The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of mediator. *Molecular and Cellular Biology*, 29(3), 650–661. <https://doi.org/10.1128/MCB.00993-08>
- Krawchuk, D., Honma-Yamanaka, N., Anani, S., & Yamanaka, Y. (2013). FGF4 is a limiting factor controlling the proportions of primitive endoderm and epiblast in the ICM of the mouse blastocyst. *Developmental Biology*, 384(1), 65–71. <https://doi.org/10.1016/j.ydbio.2013.09.023>
- Kremer, S. B., Kim, S., Jeon, J. O., Moustafa, Y. W., Chen, A., Zhao, J., & Gross, D. S. (2012). Role of Mediator in Regulating Pol II Elongation and Nucleosome Displacement in *Saccharomyces cerevisiae*. *Genetics*, 191(1), 95–106. <https://doi.org/10.1534/GENETICS.111.135806>
- Lackner, A., Sehlke, R., Garmhausen, M., Giuseppe Stirparo, G., Huth, M., Titz-Teixeira, F., Lelij, P., Ramesmayer, J., Thomas, H. F., Ralser, M., Santini, L., Galimberti, E., Sarov, M., Stewart, A. F., Smith, A., Beyer, A., & Leeb, M. (2021). Cooperative genetic networks drive embryonic stem cell transition from naïve to formative pluripotency. *The EMBO Journal*, 40(8). <https://doi.org/10.15252/embj.2020105776>
- Lai, F., Orom, U. A., Cesaroni, M., Beringer, M., Taatjes, D. J., Blobel, G. A., & Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 2013 494:7438, 494(7438), 497–501. <https://doi.org/10.1038/nature11884>

- Lake, D., Corrêa, S. A. L., & Müller, J. (2016). Negative feedback regulation of the ERK1/2 MAPK pathway. *Cellular and Molecular Life Sciences*, 73(23), 4397. <https://doi.org/10.1007/S00018-016-2297-8>
- Larsson, A. J. M., Johnsson, P., Hagemann-Jensen, M., Hartmanis, L., Faridani, O. R., Reinius, B., Segerstolpe, Å., Rivera, C. M., Ren, B., & Sandberg, R. (2019). Genomic encoding of transcriptional burst kinetics. *Nature*, 565(7738), 251–254. <https://doi.org/10.1038/s41586-018-0836-1>
- Li, C., Cesbron, F., Oehler, M., Brunner, M., & Höfer, T. (2018). Frequency Modulation of Transcriptional Bursting Enables Sensitive and Rapid Gene Regulation. *Cell Systems*, 6(4), 409–423. <https://doi.org/10.1016/j.cels.2018.01.012>
- Li, M., Yu, J. S. L., Tilgner, K., Ong, S. H., Koike-Yusa, H., & Yusa, K. (2018). Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution. *Cell Reports*, 24(2), 489–502. <https://doi.org/10.1016/j.celrep.2018.06.027>
- Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., Irizarry, R. A., Liu, J. S., Brown, M., & Liu, X. S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology*, 15(554), 1–12. <https://doi.org/10.1186/s13059-014-0554-4>
- Li, Y. C., Chao, T. C., Kim, H. J., Cholko, T., Chen, S. F., Li, G., Snyder, L., Nakanishi, K., Chang, C. E., Murakami, K., Garcia, B. A., Boyer, T. G., & Tsai, K. L. (2021). Structure and noncanonical Cdk8 activation mechanism within an Argonaute-containing Mediator kinase module. *Science Advances*, 7(3). <https://doi.org/10.1126/SCIADV.ABD4484>
- Liu, G. Y., & Sabatini, D. M. (2020). mTOR at the nexus of nutrition, growth, ageing and disease. *Nature Reviews Molecular Cell Biology*, 21(4), 183–203. <https://doi.org/10.1038/s41580-019-0199-y>
- Liu, Y., Yu, C., Daley, T. P., Wang, F., Cao, W. S., Bhate, S., Lin, X., Still, C., Liu, H., Zhao, D., Wang, H., Xie, X. S., Ding, S., Wong, W. H., Wernig, M., & Qi, L. S. (2018). CRISPR Activation Screens Systematically Identify Factors that Drive Neuronal Fate and Reprogramming. *Cell Stem Cell*, 23(5), 758–771. <https://doi.org/10.1016/j.STEM.2018.09.003>
- Long, X., Lin, Y., Ortiz-Vega, S., Yonezawa, K., & Avruch, J. (2005). Rheb Binds and Regulates the mTOR Kinase. *Current Biology*, 15(8), 702–713. <https://doi.org/10.1016/J.CUB.2005.02.053>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Luyties, O., & Taatjes, D. J. (2022). The Mediator kinase module: an interface between cell signaling and transcription. *Trends in Biochemical Sciences*, 47(4), 314–327. <https://doi.org/10.1016/j.tibs.2022.01.002>
- Lynch, C. J., Bernad, R., Martínez-Val, A., Shahbazi, M. N., Nóbrega-Pereira, S., Calvo, I., Blanco-Aparicio, C., Tarantino, C., Garreta, E., Richart-Ginés, L., Alcazar, N., Graña-Castro, O., Gómez-Lopez, G., Aksoy, I., Muñoz-Martín, M., Martínez, S., Ortega, S., Prieto, S., Simboeck, E., ... Serrano, M. (2020). Global hyperactivation of enhancers stabilizes human and mouse naive pluripotency through inhibition of CDK8/19 Mediator kinases. *Nature Cell Biology*, 22(10), 1223–1238. <https://doi.org/10.1038/s41556-020-0573-1>
- Ma, L., Chen, Z., Erdjument-Bromage, H., Tempst, P., & Pandolfi, P. P. (2005). Phosphorylation and functional inactivation of TSC2 by Erk implications for tuberous sclerosis and cancer pathogenesis. *Cell*, 121(2), 179–193. <https://doi.org/10.1016/j.cell.2005.02.031>
- Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., & Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell*, 11(4), 491–504. <https://doi.org/10.1016/J.STEM.2012.06.008>

- Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences*, 78(12), 7634–7638. <https://doi.org/10.1073/PNAS.78.12.7634>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), 10–12. <https://doi.org/10.14806/EJ.17.1.200>
- Matsuo, I., & Kimura-Yoshida, C. (2013). Extracellular modulation of Fibroblast Growth Factor signaling through heparan sulfate proteoglycans in mammalian development. *Current Opinion in Genetics & Development*, 23(4), 399–407. <https://doi.org/10.1016/J.GDE.2013.02.004>
- Mayer, C., Zhao, J., Yuan, X., & Grummt, I. (2004). mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. *Genes & Development*, 18(4), 423–434. <https://doi.org/10.1101/GAD.285504>
- McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. <http://arxiv.org/abs/1802.03426>
- Molotkov, A., Mazot, P., Brewer, J. R., Cinalli, R. M., & Soriano, P. (2017). Distinct Requirements for FGFR1 and FGFR2 in Primitive Endoderm Development and Exit from Pluripotency. *Developmental Cell*, 41(5), 511–526. <https://doi.org/10.1016/j.devcel.2017.05.004>
- Morgani, S. M., Canham, M. A., Nichols, J., Sharov, A. A., Migueles, R. P., Ko, M. S. H., & Brickman, J. M. (2013). Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Reports*, 3(6), 1945–1957. <https://doi.org/10.1016/J.CELREP.2013.04.034>
- Morgani, S. M., Nichols, J., & Hadjantonakis, A. K. (2017). The many faces of Pluripotency: in vitro adaptations of a continuum of in vivo states. *BMC Developmental Biology*, 17(1), 1–20. <https://doi.org/10.1186/S12861-017-0150-4>
- Morgani, S. M., Saiz, N., Garg, V., Raina, D., Simon, C. S., Kang, M., Arias, A. M., Nichols, J., Schröter, C., & Hadjantonakis, A. K. (2018). A Sprouty4 reporter to monitor FGF/ERK signaling activity in ESCs and mice. *Developmental Biology*, 441(1), 104–126. <https://doi.org/10.1016/j.ydbio.2018.06.017>
- Mylona, A., Theillet, F. X., Foster, C., Cheng, T. M., Miralles, F., Bates, P. A., Selenko, P., & Treisman, R. (2016). Opposing effects of Elk-1 multisite phosphorylation shape its response to ERK activation. *Science*, 354(6309), 233–237. https://doi.org/10.1126/SCIENCE.AAD1872/SUPPL_FILE/MYLONA.SM.PDF
- Näär, A. M., Taatjes, D. J., Zhai, W., Nogales, E., & Tjian, R. (2002). Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes & Development*, 16(11), 1339–1344. <https://doi.org/10.1101/GAD.987602>
- Neagu, A., van Genderen, E., Escudero, I., Verwegen, L., Kurek, D., Lehmann, J., Stel, J., Dirks, R. A. M., van Mierlo, G., Maas, A., Eleveld, C., Ge, Y., den Dekker, A. T., Brouwer, R. W. W., van IJcken, W. F. J., Modic, M., Drukker, M., Jansen, J. H., Rivron, N. C., ... ten Berge, D. (2020). In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states. *Nature Cell Biology*, 22(5), 534–545. <https://doi.org/10.1038/S41556-020-0508-X>
- Nguyen, L. Van, Laval, J.-P., Chainais, P., Iop, A., Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 1–12. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Nichols, J., Evans, E. P., & Smith, A. G. (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development*, 110(4), 1341–1348. <https://doi.org/10.1242/DEV.110.4.1341>
- Nichols, J., Silva, J., Roode, M., & Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development*, 136(19), 3215–3222. <https://doi.org/10.1242/dev.038893>

- Niwa, H., Ogawa, K., Shimosato, D., & Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*, *460*(7251), 118–122. <https://doi.org/10.1038/nature08113>
- Nowotschin, S., Costello, I., Piliszek, A., Kwon, G. S., Mao, C. an, Klein, W. H., Robertson, E. J., & Hadjantonakis, A. K. (2013). The T-box transcription factor Eomesodermin is essential for AVE induction in the mouse embryo. *Genes & Development*, *27*(9), 997–1002. <https://doi.org/10.1101/GAD.215152.113>
- Nusse, R. (2012). Wnt Signaling. *Cold Spring Harbor Perspectives in Biology*, *4*(5), 1–4. <https://doi.org/10.1101/CSHPERSPECT.A011163>
- Ochiai, H., Hayashi, T., Umeda, M., Yoshimura, M., Harada, A., Shimizu, Y., Nakano, K., Saitoh, N., Liu, Z., Yamamoto, T., Okamura, T., Ohkawa, Y., Kimura, H., & Nikaido, I. (2020). Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Science Advances*, *6*(25), 6699–6716. <https://doi.org/10.1126/sciadv.aaz6699>
- Ochiai, H., Sugawara, T., Sakuma, T., & Yamamoto, T. (2014). Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells. *Scientific Reports*, *4*, 1–9. <https://doi.org/10.1038/srep07125>
- Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oleå, A. K., Araúzo-Bravo, M. J., Saitou, M., Hadjantonakis, A. K., & Hiiragi, T. (2013). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nature Cell Biology*, *16*(1), 27–37. <https://doi.org/10.1038/ncb2881>
- Ornitz, D. M., & Itoh, N. (2015). The fibroblast growth factor signaling pathway. *Wiley Interdisciplinary Reviews: Developmental Biology*, *4*(3), 215–266. <https://doi.org/10.1002/wdev.176>
- Peccoud, J., & Ycard, B. (1995). Markovian Modelling of Gene Product Synthesis. *Theoretical Population Biology*, *48*, 222–234. <https://doi.org/https://doi.org/10.1006/tpbi.1995.1027>
- Pereira, L., Yi, F., & Merrill, B. J. (2006). Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Molecular and Cellular Biology*, *26*(20), 7479–7491. <https://doi.org/10.1128/MCB.00368-06>
- Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., & Hadjantonakis, A. K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development*, *135*(18), 3081–3091. <https://doi.org/10.1242/DEV.021519>
- Raina, D., Bahadori, A., Stanoev, A., Protzek, M., Koseska, A., & Schröter, C. (2021). Cell-cell communication through FGF4 generates and maintains robust proportions of differentiated cell types in embryonic stem cells. *Development*, *148*(21), 1–31. <https://doi.org/10.1242/DEV.199926>
- Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., & Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS Biology*, *4*(10), 1707–1719. <https://doi.org/10.1371/journal.pbio.0040309>
- Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., & Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods*, *5*(10), 877–879. <https://doi.org/10.1038/NMETH.1253>
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. <https://doi.org/10.1038/NPROT.2013.143>
- Rathjen, J. (2014). The States of Pluripotency: Pluripotent Lineage Development in the Embryo and in the Dish. *ISRN Stem Cells*, *2014*, 1–19. <https://doi.org/10.1155/2014/208067>
- Replogle, J. M., Norman, T. M., Xu, A., Hussmann, J. A., Chen, J., Cogan, J. Z., Meer, E. J., Terry, J. M., Riordan, D. P., Srinivas, N., Fiddes, I. T., Arthur, J. G., Alvarado, L. J., Pfeiffer, K. A., Mikkelsen, T. S., Weissman, J. S., & Adamson, B. (2020). Combinatorial single-cell CRISPR screens by direct

- guide RNA capture and targeted sequencing. *Nature Biotechnology*, 38(8), 954–961. <https://doi.org/10.1038/s41587-020-0470-y>
- Ring, D. B., Johnson, K. W., Henriksen, E. J., Nuss, J. M., Goff, D., Kinnick, T. R., Ma, S. T., Reeder, J. W., Samuels, I., Slabiak, T., Wagman, A. S., Hammond, M. E. W., & Harrison, S. D. (2003). Selective Glycogen Synthase Kinase 3 Inhibitors Potentiate Insulin Activation of Glucose Transport and Utilization In Vitro and In Vivo. *Diabetes*, 52(3), 588–595. <https://doi.org/10.2337/DIABETES.52.3.588>
- Rocha, P. P., Scholze, M., Bleiß, W., & Schrewe, H. (2010). Med12 is essential for early mouse development and for canonical Wnt and Wnt/PCP signaling. *Development*, 137(16), 2723–2731. <https://doi.org/10.1242/dev.053660>
- Rosales-Alvarez, R. E., Rettkowski, J., Herman, J. S., Dumbović, G., Cabezas-Wallscheid, N., & Grün, D. (2023). VarID2 quantifies gene expression noise dynamics and unveils functional heterogeneity of ageing hematopoietic stem cells. *Genome Biology*, 24(1). <https://doi.org/10.1186/s13059-023-02974-1>
- Saiz, N., & Plusa, B. (2013). Early cell fate decisions in the mouse embryo. *Reproduction*, 145(3), R65–R80. <https://doi.org/10.1530/REP-12-0381>
- Schindelin, J., Arganda-carreras, I., Frise, E., Kaynig, V., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J., White, D. J., Hartenstein, V., Tomancak, P., & Cardona, A. (2012). Fiji - an Open Source platform for biological image analysis. *Nature Methods*, 9(7), 676–682. <https://doi.org/10.1038/nmeth.2019>
- Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell detection with star-convex polygons. *Lecture Notes in Computer Science*, 11071 LNCS, 265–273. https://doi.org/https://doi.org/10.1007/978-3-030-00934-2_30
- Schrode, N., Saiz, N., Di Talia, S., & Hadjantonakis, A. K. (2014). GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. *Developmental Cell*, 29(4), 454–467. <https://doi.org/10.1016/J.DEVCEL.2014.04.011>
- Schröter, C., Rué, P., Mackenzie, J. P., & Arias, A. M. (2015). FGF/MAPK signaling sets the switching threshold of a bistable circuit controlling cell fate decisions in embryonic stem cells. *Development*, 142(24), 4205–4216. <https://doi.org/10.1242/dev.127530>
- Schumacher, S., Fernkorn, M., Marten, M., Kim, Y. S., Bedzhov, I., & Schröter, C. (2023). Tissue-intrinsic Wnt signals antagonize Nodal-driven AVE differentiation. *BioRxiv*, 1–34. <https://doi.org/10.1101/2023.05.19.541432>
- Schwabe, D., Formichetti, S., Junker, J. P., Falcke, M., & Rajewsky, N. (2020). The transcriptome dynamics of single cells during the cell cycle. *Molecular Systems Biology*, 16(11). <https://doi.org/10.15252/MSB.20209946>
- Seruggia, D., Oti, M., Tripathi, P., Canver, M. C., LeBlanc, L., Di Giammartino, D. C., Bullen, M. J., Nefzger, C. M., Sun, Y. B. Y., Farouni, R., Polo, J. M., Pinello, L., Apostolou, E., Kim, J., Orkin, S. H., & Das, P. P. (2019). TAF5L and TAF6L Maintain Self-Renewal of Embryonic Stem Cells via the MYC Regulatory Network. *Molecular Cell*, 74(6), 1148–1163. <https://doi.org/10.1016/j.molcel.2019.03.025>
- Shao, R., Kumar, B., Lidschreiber, K., Lidschreiber, M., Cramer, P., & Elsässer, S. J. (2022). Distinct transcription kinetics of pluripotent cell states. *Molecular Systems Biology*, 18(1), 1–19. https://doi.org/10.15252/MSB.202110407/SUPPL_FILE/MSB202110407-SUP-0003-TABLEEV2.XLSX
- Sharrocks, A. D. (2001). The ETS-domain transcription factor family. *Nature Reviews Molecular Cell Biology*, 2(11), 827–837. <https://doi.org/10.1038/35099076>
- Simon, C. S., Hadjantonakis, A. K., & Schröter, C. (2018). Making lineage decisions with biological noise: Lessons from the early mouse embryo. In *Wiley Interdisciplinary Reviews: Developmental Biology* (Vol. 7, Issue 4). John Wiley and Sons Inc. <https://doi.org/10.1002/wdev.319>

- Smith, A. (2017). Formative pluripotency: The executive phase in a developmental continuum. *Development*, *144*(3), 365–373. <https://doi.org/10.1242/dev.142679>
- Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology*, *19*(4), 262–274. <https://doi.org/10.1038/nrm.2017.115>
- Stavridis, M. P., Simon Lunn, J., Collins, B. J., & Storey, K. G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development*, *134*(16), 2889–2894. <https://doi.org/10.1242/DEV.02858>
- Steklov, M., Pandolfi, S., Baietti, M. F., Batiuk, A., Carai, P., Najm, P., Zhang, M., Jang, H., Renzi, F., Cai, Y., Abbasi Asbagh, L., Pastor, T., De Troyer, M., Simicek, M., Radaelli, E., Brems, H., Legius, E., Tavernier, J., Gevaert, K., ... Sablina, A. A. (2018). Mutations in LZTR1 drive human disease by dysregulating RAS ubiquitination. *Science*, *362*(6419), 1177–1182. <https://doi.org/10.1126/SCIENCE.AAP7607>
- Stevens, J. L., Cantin, G. T., Wang, G., Shevchenko, A., Shevchenko, A., & Berk, A. J. (2002). Transcription control by E1A and MAP kinase pathway via Sur2 mediator subunit. *Science*, *296*(5568), 755–758. <https://doi.org/10.1126/SCIENCE.1068943>
- Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., & Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophoblast in the mouse blastocyst. *Development*, *132*(9), 2093–2102. <https://doi.org/10.1242/DEV.01801>
- Svejstrup, J. Q. (2007). Elongator complex: how many roles does it play? *Current Opinion in Cell Biology*, *19*(3), 331–336. <https://doi.org/10.1016/j.ceb.2007.04.005>
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., & Von Mering, C. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, *43*(Database issue), D447–D452. <https://doi.org/10.1093/NAR/GKU1003>
- Tang, M., Kaymaz, Y., Logeman, B. L., Eichhorn, S., Liang, Z. S., Dulac, C., & Sackton, T. B. (2021). Evaluating single-cell cluster stability using the Jaccard similarity index. *Bioinformatics*, *37*(15), 2212–2214. <https://doi.org/10.1093/BIOINFORMATICS/BTAA956>
- Tarazi, S., Aguilera-Castrejon, A., Joubran, C., Ghanem, N., Ashouokhi, S., Roncato, F., Wildschutz, E., Haddad, M., Oldak, B., Gomez-Cesar, E., Livnat, N., Viukov, S., Lokshtanov, D., Naveh-Tassa, S., Rose, M., Hanna, S., Raanan, C., Brenner, O., Kedmi, M., ... Hanna, J. H. (2022). Post-gastrulation synthetic embryos generated ex utero from mouse naive ESCs. *Cell*, *185*, 3290–3306. <https://doi.org/10.1016/J.CELL.2022.07.028>
- Tee, W. W., Shen, S. S., Oksuz, O., Narendra, V., & Reinberg, D. (2014). Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. *Cell*, *156*(4), 678–690. <https://doi.org/10.1016/j.cell.2014.01.009>
- Tepekoy, F., Akkoyunlu, G., & Demir, R. (2015). The role of Wnt signaling members in the uterus and embryo during pre-implantation and implantation. *Journal of Assisted Reproduction and Genetics*, *32*(3), 337–346. <https://doi.org/10.1007/S10815-014-0409-7>
- ter Huurne, M., Peng, T., Yi, G., van Mierlo, G., Marks, H., & Stunnenberg, H. G. (2020). Critical Role for P53 in Regulating the Cell Cycle of Ground State Embryonic Stem Cells. *Stem Cell Reports*, *14*(2), 175–183. <https://doi.org/10.1016/J.STEMCR.2020.01.001>
- Tsai, K. L., Sato, S., Tomomori-Sato, C., Conaway, R. C., Conaway, J. W., & Asturias, F. J. (2013). A conserved Mediator–CDK8 kinase module association regulates Mediator–RNA polymerase II interaction. *Nature Structural & Molecular Biology*, *20*(5), 611–619. <https://doi.org/10.1038/nsmb.2549>
- Tutter, A. V., Kowalski, M. P., Baltus, G. A., Iourgenko, V., Labow, M., Li, E., & Kadam, S. (2009). Role for Med12 in regulation of Nanog and Nanog target genes. *Journal of Biological Chemistry*, *284*(6), 3709–3718. <https://doi.org/10.1074/jbc.M805677200>

- Ünal, E. B., Uhlitz, F., & Blüthgen, N. (2017). A compendium of ERK targets. *FEBS Letters*, *591*(17), 2607–2615. <https://doi.org/10.1002/1873-3468.12740>
- Villegas, F., Lehalle, D., Mayer, D., Rittirsch, M., Stadler, M. B., Zinner, M., Olivieri, D., Vabres, P., Duplomb-Jego, L., De Bont, E. S. J. M., Duffourd, Y., Duijkers, F., Avila, M., Geneviève, D., Houcinat, N., Jouan, T., Kuentz, P., Lichtenbelt, K. D., Thauvin-Robinet, C., ... Betschinger, J. (2019). Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3. *Cell Stem Cell*, *24*(2), 257–270. <https://doi.org/10.1016/j.STEM.2018.11.021>
- Vu, T. N., Wills, Q. F., Kalari, K. R., Niu, N., Wang, L., Rantalainen, M., & Pawitan, Y. (2016). Beta-Poisson model for single-cell RNA-seq data analyses. *Bioinformatics*, *32*(14), 2128–2135. <https://doi.org/10.1093/bioinformatics/btw202>
- Wamaitha, S. E., del Valle, I., Cho, L. T. Y., Wei, Y., Fogarty, N. M. E., Blakeley, P., Sherwood, R. I., Ji, H., & Niakan, K. K. (2015). Gata6 potently initiates reprogramming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes & Development*, *29*(12), 1239–1255. <https://doi.org/10.1101/GAD.257071.114>
- Wernet, M. F., Mazzoni, E. O., Çelik, A., Duncan, D. M., Duncan, I., & Desplan, C. (2006). Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature*, *440*(7081), 174–180. <https://doi.org/10.1038/nature04615>
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I., & Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, *153*(2), 307–319. <https://doi.org/10.1016/j.cell.2013.03.035>
- Xenopoulos, P., Kang, M., Puliafito, A., DiTalia, S., & Hadjantonakis, A. K. (2015). Heterogeneities in Nanog Expression Drive Stable Commitment to Pluripotency in the Mouse Blastocyst. *Cell Reports*, *10*(9), 1508–1520. <https://doi.org/10.1016/j.CELREP.2015.02.010>
- Yamanaka, Y., Lanner, F., & Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development*, *137*(5), 715–724. <https://doi.org/10.1242/dev.043471>
- Yang, B., Kuang, J., Wu, C., Zhou, W., Zhu, S., Jiang, H., Zhai, Z., Wu, Y., Peng, J., Liu, N., Hu, H., Ide, N. M., Chen, R., Zhao, M., & Zhu, P. (2020). Screening Genes Promoting Exit from Naive Pluripotency Based on Genome-Scale CRISPR-Cas9 Knockout. *Stem Cells International*, *2020*, 1–15. <https://doi.org/10.1155/2020/8483035>
- Yang, S. H., Kalkan, T., Morrisroe, C., Smith, A., & Sharrocks, A. D. (2012). A Genome-Wide RNAi Screen Reveals MAP Kinase Phosphatases as Key ERK Pathway Regulators during Embryonic Stem Cell Differentiation. *PLOS Genetics*, *8*(12), 1–15. <https://doi.org/10.1371/JOURNAL.PGEN.1003112>
- Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., & Ornitz, D. M. (1991). Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell*, *64*(4), 841–848. [https://doi.org/10.1016/0092-8674\(91\)90512-W](https://doi.org/10.1016/0092-8674(91)90512-W)
- Yin, J. W., & Wang, G. (2014). The Mediator complex: A master coordinator of transcription and cell lineage development. *Development*, *141*(5), 977–987. <https://doi.org/10.1242/dev.098392>
- Ying, Q. L., Stavridis, M., Griffiths, D., Li, M., & Smith, A. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nature Biotechnology*, *21*(2), 183–186. <https://doi.org/10.1038/nbt780>
- Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., & Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature*, *453*(7194), 519–523. <https://doi.org/10.1038/nature06968>
- Zamudio, A. V., Dall'Agnesse, A., Henninger, J. E., Manteiga, J. C., Afeyan, L. K., Hannett, N. M., Coffey, E. L., Li, C. H., Oksuz, O., Sabari, B. R., Boija, A., Klein, I. A., Hawken, S. W., Spille, J. H., Decker, T. M., Cisse, I. I., Abraham, B. J., Lee, T. I., Taatjes, D. J., ... Young, R. A. (2019). Mediator

- Condensates Localize Signaling Factors to Key Cell Identity Genes. *Molecular Cell*, 76(5), 1–14. <https://doi.org/10.1016/j.molcel.2019.08.016>
- Zhang, X., Zhao, X., Li, G., Zhang, M., Xing, P., Li, Z., Chen, B., Yang, H., & Wu, Z. (2021). Establishment of Etv5 gene knockout mice as a recipient model for spermatogonial stem cell transplantation. *Biology Open*, 10(1). <https://doi.org/10.1242/BIO.056804>
- Zhao, M., Yang, X., Fu, Y., Wang, H., Ning, Y., Yan, J., Chen, Y. G., & Wang, G. (2013). Mediator MED15 modulates transforming growth factor beta (TGF β)/Smad signaling and breast cancer cell metastasis. *Journal of Molecular Cell Biology*, 5(1), 57–60. <https://doi.org/10.1093/JMCB/MJS054>
- Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., ... Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, 8(1), 1–12. <https://doi.org/10.1038/ncomms14049>

7 Appendix

7.1 List of used abbreviations

Abbreviation	Meaning
AKT	Protein kinase B
au	arbitrary unit
AVE	Anterior Visceral Endoderm
CRISPRa	Clustered Regularly Interspaced Short Palindromic Repeats activation
CTD	carboxyl-terminal domain (of PolII)
Epi	Epiblast
ERK	Extracellular signal-regulated kinase
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
GDP	guanosine diphosphate
gRNA	guide RNA
GTP	guanosine triphosphate
HCR	Hybridization Chain Reaction
ICM	Inner Cell Mass
JAK/STAT	Janus kinases (JAK) - signal transducer and activator of transcription proteins (STAT)
LR PCR	Long Range Polymerase Chain Reaction
mRNA	messenger ribonucleic acid
mTOR	mammalian Target Of Rapamycin
PIC	preinitiation complex
PLC	Phosphoinositide phospholipase C
PolII	RNA polymerase II
ppERK	phospho-ERK
PrE	Primitive Endoderm
RNAi	RNA interference
scRNAseq	Single-cell RNA sequencing
siRNA	Small interfering RNA
TE	Trophectoderm
TF	Transcription factor
TPM	Transcript per kilobase million
Wnt	Wingless iNTEgrated
wt	wild type

7.2 List of Figures

Figure 1: Early mouse embryonic development.....	3
Figure 2: mESCs recapitulate symmetry breaking in early development.	7
Figure 3: Schematic of FGF-signaling cascade in mouse embryonic stem cells.....	10
Figure 4: Mediator links enhancers and promoters to regulate transcription.	14
Figure 5: Telegraph model of transcriptional bursting leading to differently shaped mRNA count distributions.....	17
Figure 6: FGF4 stimulation induces strong gene expression in mESCs.....	36
Figure 7: FGF-stimulation leads to increased gene expression variability in mESCs.....	38
Figure 8: Reduced transcriptional burst frequency and increased cell-to-cell variation of transcription of FGF target genes.	40
Figure 9: <i>Spry4</i> ^{H2B-Venus/+} reporter cells as a readout for signaling-induced transcriptional activity.	41
Figure 10: Schematic of a genome-wide CRISPR knockout screen.	42
Figure 11: Enrichment of gRNAs and corresponding genes.	44
Figure 12: Genome-wide CRISPR knockout screen reveals positive regulators of <i>Sprouty4</i> expression.....	47
Figure 13: Genome-wide CRISPR knockout screen reveals negative regulators of <i>Sprouty4</i> expression.....	49
Figure 14: Generation of <i>Med12</i> mutant cell lines.....	51
Figure 15: Differentially expressed genes between wild-type and <i>Med12</i> -mutant cells in pluripotency and differentiation.....	53
Figure 16: <i>Med12</i> mutant cell's expression footprint most similar to mTOR perturbation.	54
Figure 17: <i>Med12</i> regulates gene expression largely independent of mTOR, Wnt, and FGF-signaling.....	56
Figure 18: Limited redundancy between <i>Med12</i> and <i>Med12L</i>	57
Figure 19: Mutation of <i>Med12</i> decreases clonogenicity.	59
Figure 20: Differentiation delay in <i>Med12</i> -mutant cells.	61
Figure 21: <i>Gata6</i> -inducible line with induction level and capability to differentiate into PrE similar to <i>Gata4</i> -inducible line.	62
Figure 22: Transition between embryonic and extraembryonic identities is buffered against loss of <i>Med12</i>	64
Figure 23: Loss of <i>Med12</i> affects <i>Gata6</i> -induction dynamics.....	66

Figure 24: AVE differentiation in <i>Med12</i> -mutant cells.....	68
Figure 25: Multiplexed single-cell RNA sequencing of wild-type and <i>Med12</i> -mutant cells.	70
Figure 26: Loss of <i>Med12</i> affects pluripotency homogenously.	71
Figure 27: mRNA counts increase less during differentiation in <i>Med12</i> -mutant cells.	73
Figure 28: Limited role of <i>Med12</i> in cell type-specific gene regulation.....	75
Figure 29: Role of <i>Med12</i> in separation between cell types and transcriptional noise.....	77

7.3 List of Supplementary Figures

Supplementary Figure 1: gRNA counts at different timepoints for essential genes.....	106
Supplementary Figure 2: gRNA counts at different time points for genes inhibiting cell growth.	107
Supplementary Figure 3: gRNA counts at different time points for genes required for <i>Spry4</i> expression.....	108
Supplementary Figure 4: gRNA counts at different time points for genes inhibiting <i>Spry4</i> expression.....	109
Supplementary Figure 5: gRNA counts at different time points for Mediator subunits.	110
Supplementary Figure 6: High sequencing depth in single-cell RNA sequencing comparing FGF concentrations.	111
Supplementary Figure 7: Clear separation between signal and noise in distributions of multiplexing labels.....	111
Supplementary Figure 8: High data quality in multiplexed single cell RNA sequencing experiment.	112

7.4 List of Tables

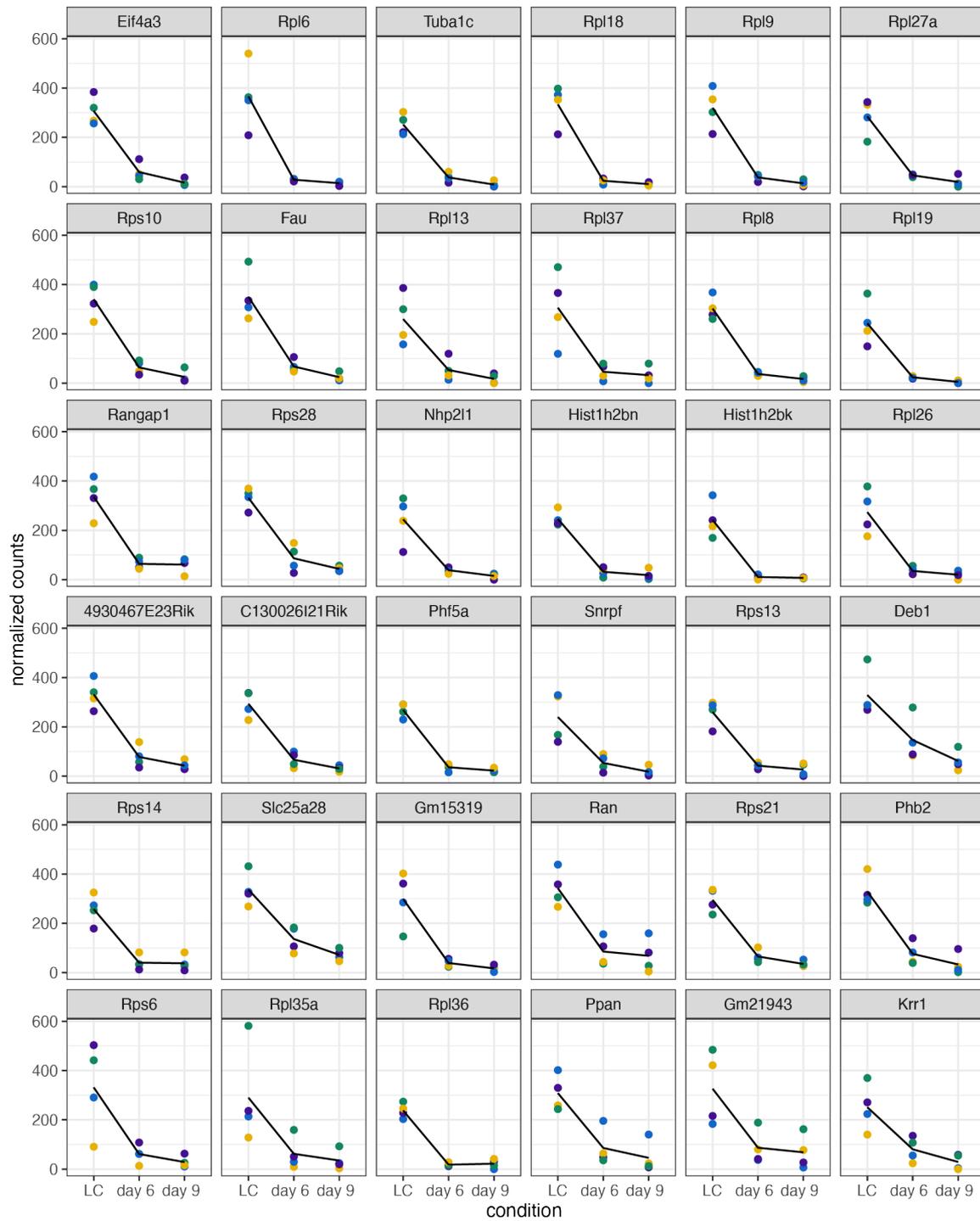
Table 1: Used primer sequences with indicated function.	20
Table 2: Used gRNAs for indicated genes.	21

7.5 List of Supplementary Tables

Supplementary Table 1: Quality metrics of filtered single-cell RNA sequencing samples.....	113
Supplementary Table 2: Quality metrics for each demultiplexed single cell RNA sequencing sample from CellRanger.	113
Supplementary Table 3: Quality metrics of filtered, demultiplexed single cell RNA sequencing samples.....	113

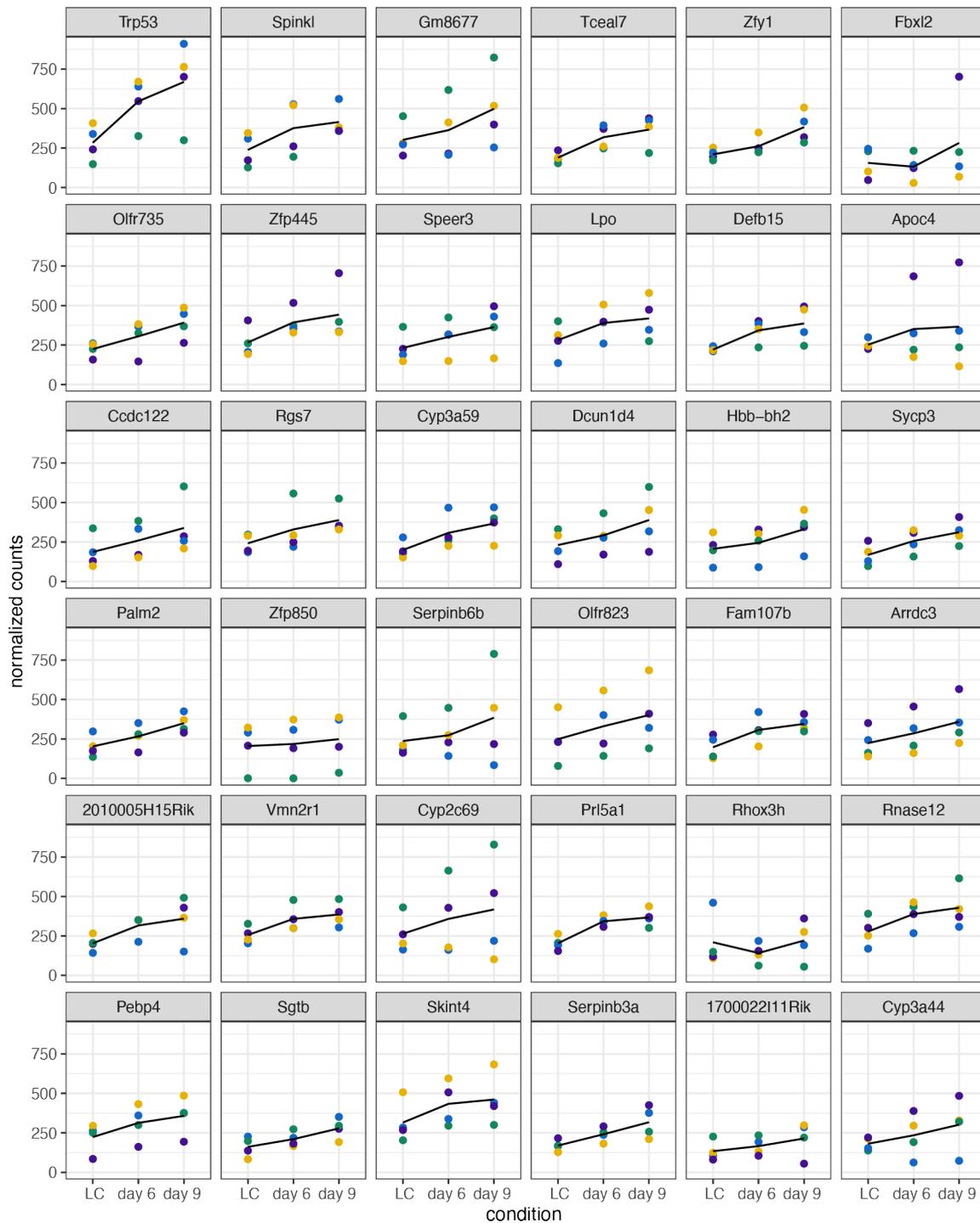
7.6 Supplementary Figures

7.6.1 gRNA representation over time during CRISPR knockout screen



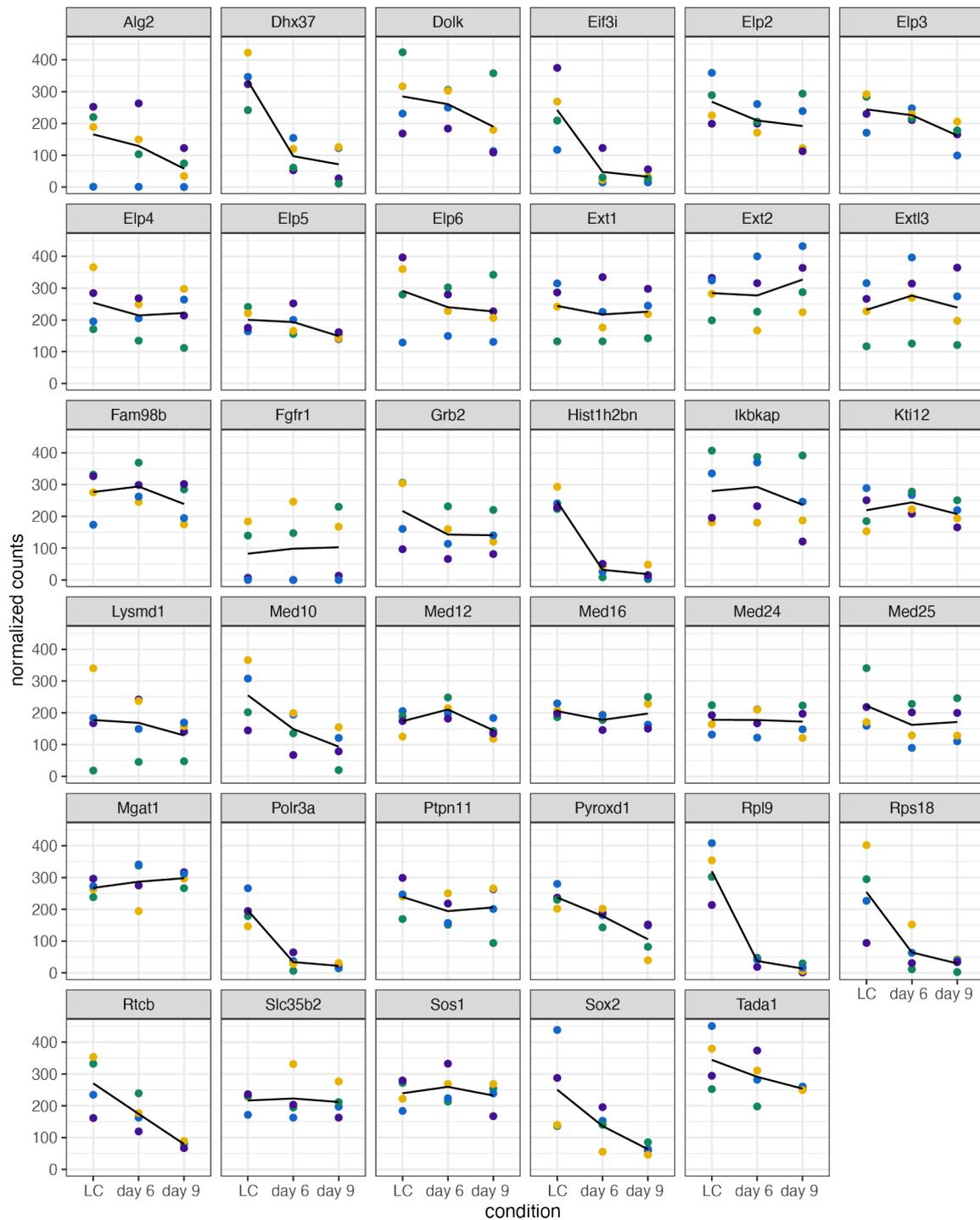
Supplementary Figure 1: gRNA counts at different timepoints for essential genes.

Essential genes defined as genes for which the representation of the corresponding gRNAs is most decreased. Shown are the 36 genes with strongest decrease between LC (library control) and non-sorted LC (library control on day 9). Colors indicate gRNA identity.



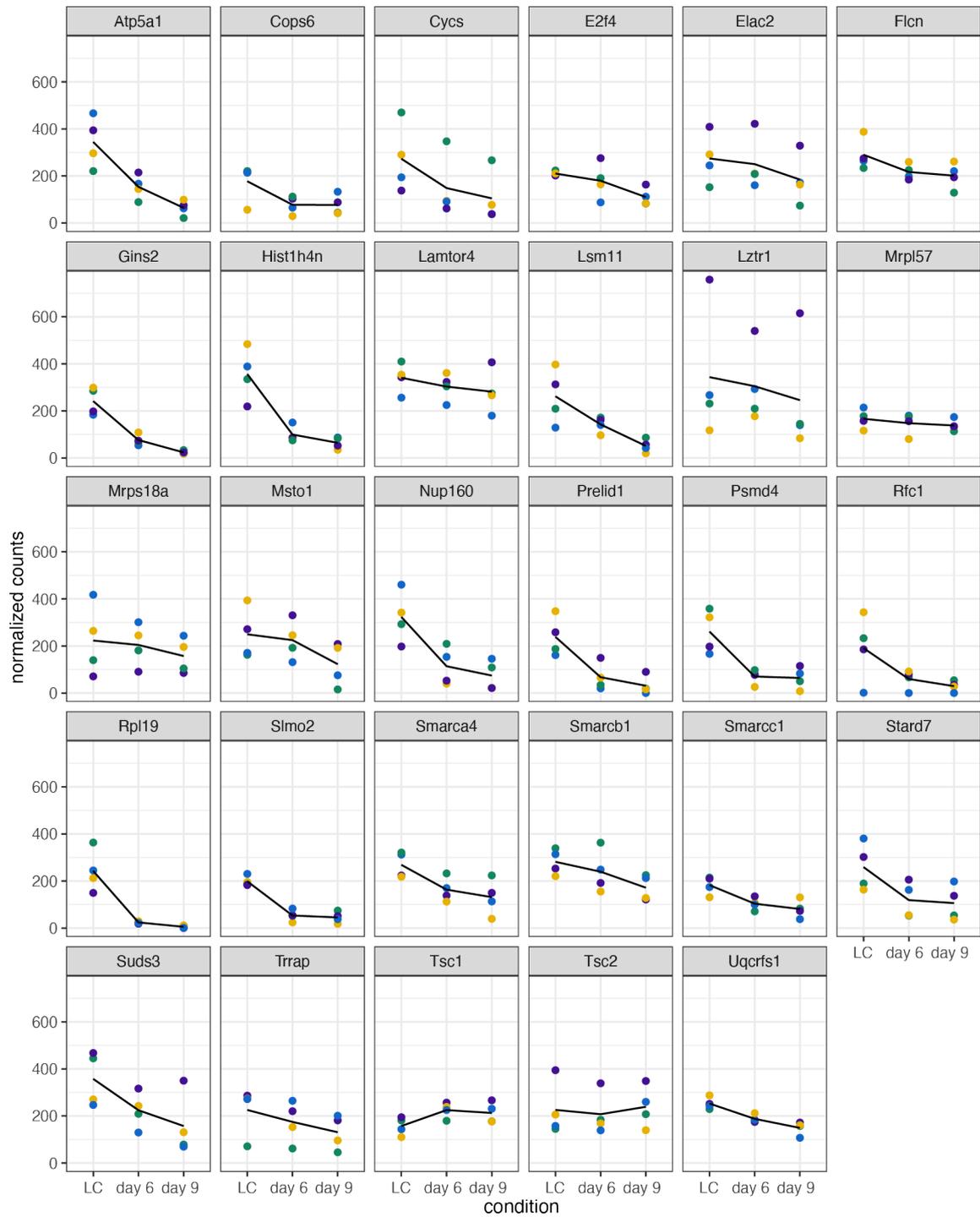
Supplementary Figure 2: gRNA counts at different time points for genes inhibiting cell growth.

Genes inhibiting cell growth are defined as genes for which the representation of the corresponding gRNAs is most increased. Shown are the 36 genes with the strongest increase between LC (library control) and non-sorted control on day 9). Colors indicate gRNA identity.

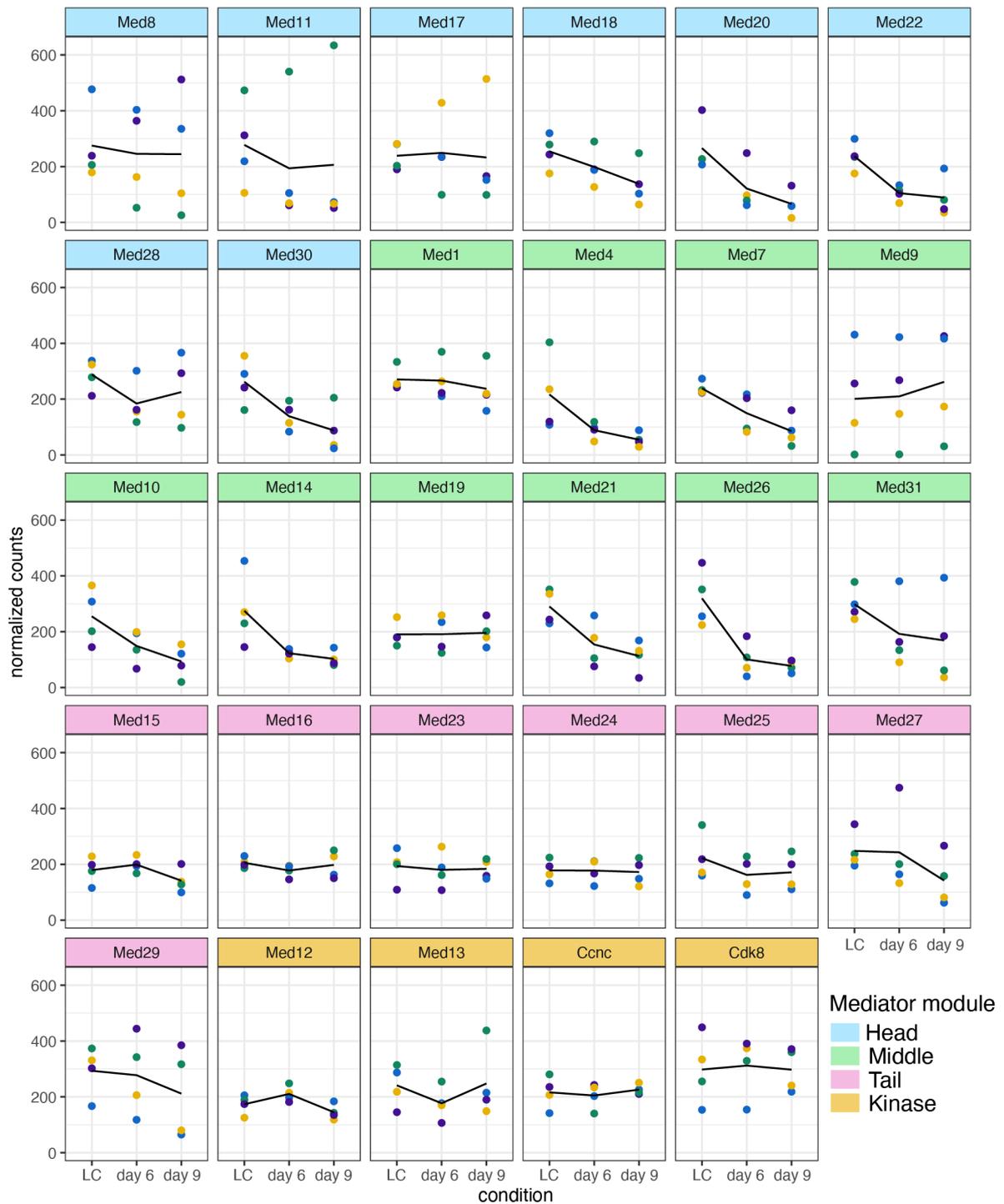


Supplementary Figure 3: gRNA counts at different time points for genes required for *Spry4* expression.

gRNA counts of corresponding genes required for efficient *Spry4* expression (selected as in Figure 12) in LC (library control) and non-sorted control on day 6 and day 9. Colors indicate gRNA identity.



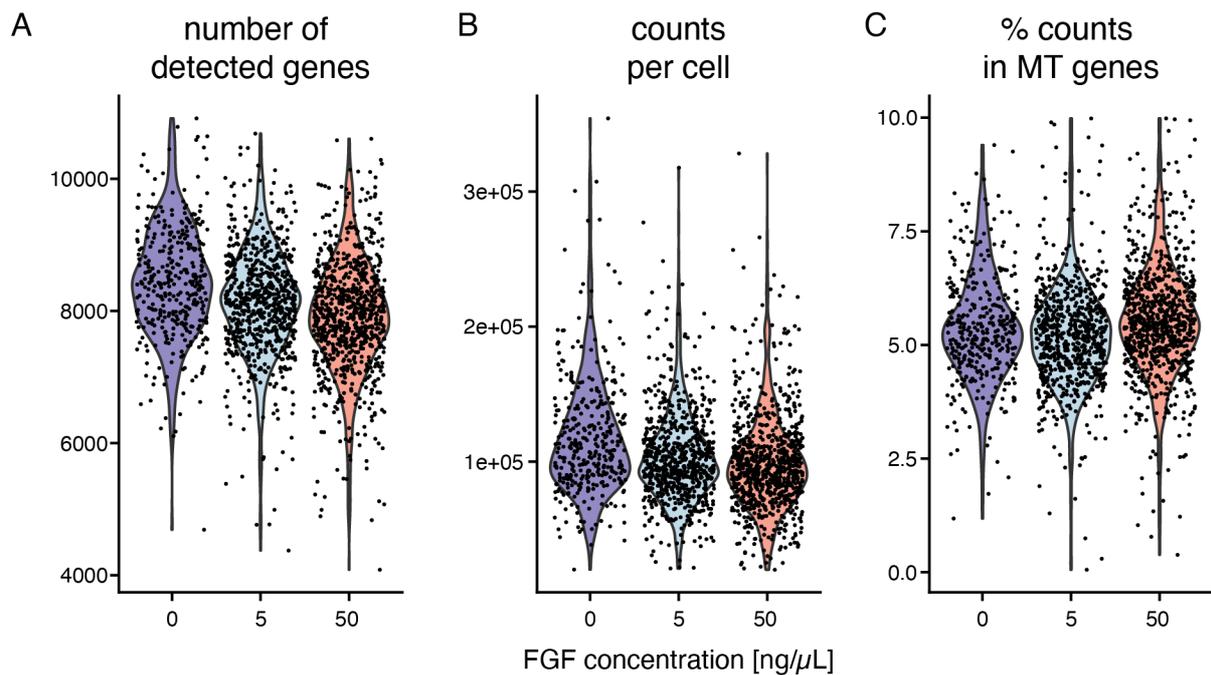
Supplementary Figure 4: gRNA counts at different time points for genes inhibiting *Spry4* expression. gRNA counts of corresponding genes inhibiting *Spry4* expression (selected as in Figure 13) in LC (library control) and non-sorted control on day 6 and day 9. Colors indicate gRNA identity.



Supplementary Figure 5: gRNA counts at different time points for Mediator subunits.

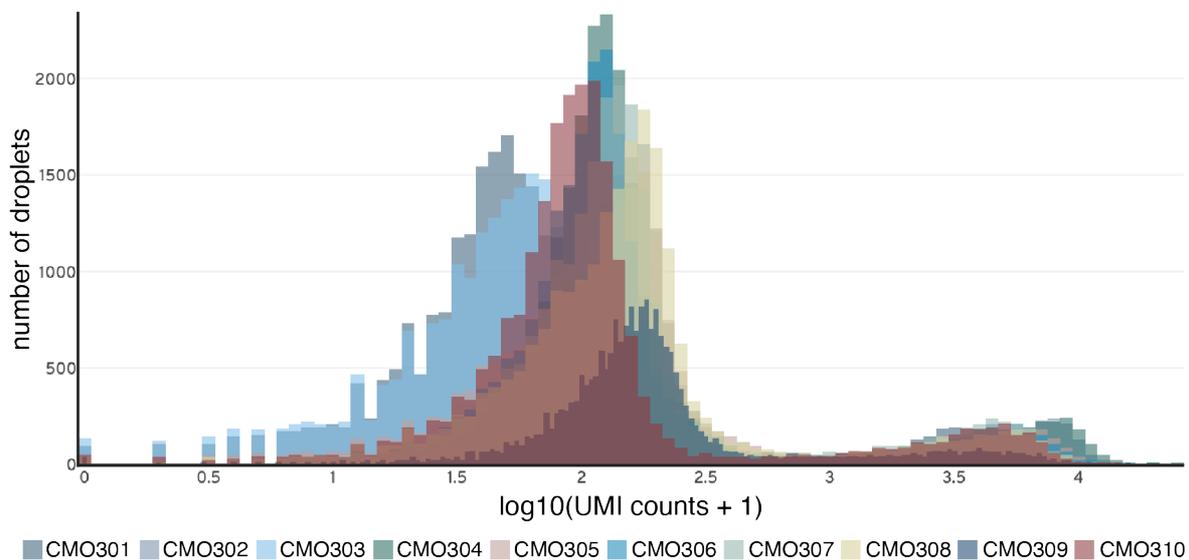
Mediator subunit-targeting gRNA counts in LC (library control) and non-sorted control on day 6 and day 9. Point colors indicate gRNA identity, the heading color indicates Mediator module of the respective subunit.

7.6.2 Single-cell RNA sequencing quality control



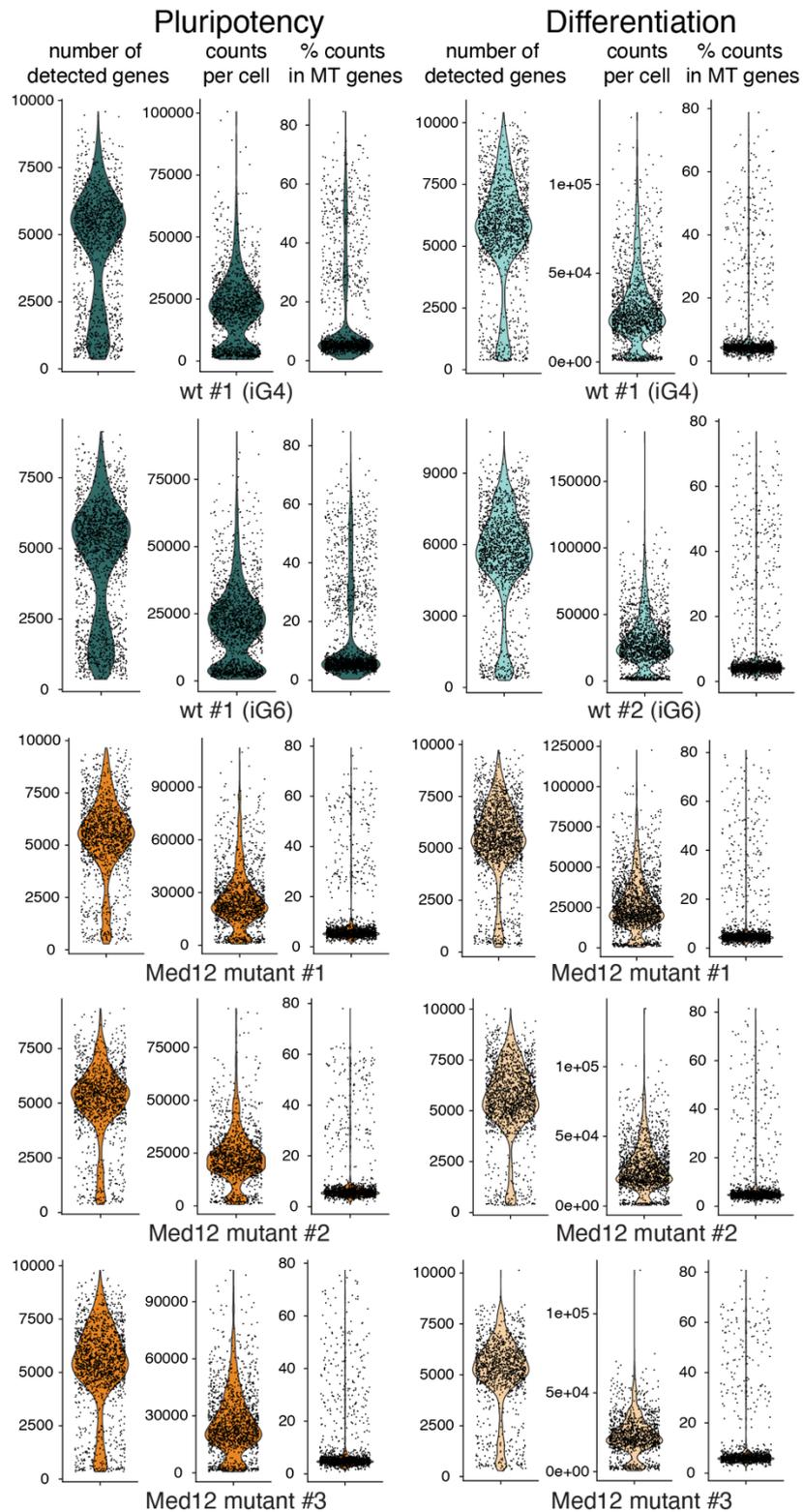
Supplementary Figure 6: High sequencing depth in single-cell RNA sequencing comparing FGF concentrations.

A - C Number of detected genes per cell (A), total counts per cell (B), and % of counts in mitochondrial (MT) genes (C) for each sample separately. Cells were filtered before plotting for a minimal number of 4000 detected genes per cell and less than 10% mitochondrial mRNA for further analysis.



Supplementary Figure 7: Clear separation between signal and noise in distributions of multiplexing labels.

Cell multiplexing oligos (CMOs) named as provided by 10x Genomics and correspond to sample as in Supplementary Table 2.



Supplementary Figure 8: High data quality in multiplexed single cell RNA sequencing experiment. Number of detected genes per cell, total counts per cell and % of counts in mitochondrial (MT) genes for each sample separately. After plotting, cells were filtered for a minimal number of 2500 detected genes per cell and less than 15% mitochondrial mRNA for further analysis.

7.7 Supplementary Tables

7.7.1 Single-cell RNA sequencing quality control

Supplementary Table 1: Quality metrics of filtered single-cell RNA sequencing samples.

Genotype	FGF Concentration	Cell number	Median detected genes	Median counts per cells	Percent counts in mitochondrial genes
<i>Fgf4</i> -mutant	0 ng/mL	361	8424	109942	5.37
<i>Fgf4</i> -mutant	5 ng/mL	713	8161	97684	5.18
<i>Fgf4</i> -mutant	50 ng/mL	808	7915	93557	5.52

Supplementary Table 2: Quality metrics for each demultiplexed single cell RNA sequencing sample from Cell Ranger.

Genotype	Conditions	CMO Name	Fraction reads in cell-associated barcodes	Cells assigned to CMO	CMO signal-to-noise ratio	Number of reads from cells called from this sample	Mapped to genome	Confidently mapped to genome	Confidently mapped to transcriptome	Confidently mapped to intronic regions	Confidently mapped to exonic regions	Confidently mapped to intergenic regions	Confidently mapped antisense
<i>Med12</i> wild type #1 (iG4)	2i + LIF	CMO 301	66.64%	9.09%	5.56	53,913,689	94.89%	90.70%	80.76%	13.99%	72.08%	4.64%	4.97%
<i>Med12</i> wild type #1 (iG4)	Dif	CMO 302	61.87%	11.53%	4.50	86,729,539	95.13%	91.32%	82.34%	13.10%	74.03%	4.19%	4.45%
<i>Med12</i> wild type #2 (iG6)	2i + LIF	CMO 303	69.31%	7.83%	5.14	67,231,902	94.90%	90.90%	80.95%	14.45%	71.98%	4.48%	5.13%
<i>Med12</i> wild type #2 (iG6)	Dif	CMO 304	65.42%	9.48%	5.95	89,291,791	95.47%	92.03%	83.71%	13.54%	74.84%	3.66%	4.30%
<i>Med12</i> mutant #1 (iG6)	2i + LIF	CMO 305	54.27%	8.71%	4.13	71,060,061	95.31%	91.64%	82.29%	14.47%	73.01%	4.16%	4.83%
<i>Med12</i> mutant #1 (iG6)	Dif	CMO 306	57.76%	10.21%	5.37	90,314,056	95.56%	92.19%	83.84%	14.58%	74.13%	3.47%	4.51%
<i>Med12</i> mutant #2 (iG6)	2i + LIF	CMO 307	52.92%	10.87%	4.94	57,515,309	95.39%	91.77%	82.54%	14.31%	73.43%	4.04%	4.84%
<i>Med12</i> mutant #2 (iG6)	Dif	CMO 308	51.49%	9.82%	4.19	81,278,964	95.62%	92.26%	84.15%	14.50%	74.30%	3.46%	4.29%
<i>Med12</i> mutant #3 (iG6)	2i + LIF	CMO 309	50.07%	11.99%	4.46	51,163,906	95.27%	91.58%	82.12%	15.17%	72.34%	4.08%	5.04%
<i>Med12</i> mutant #3 (iG6)	Dif	CMO 310	59.84%	10.49%	5.35	72,273,730	95.53%	92.14%	83.79%	14.51%	74.10%	3.53%	4.45%

Supplementary Table 3: Quality metrics of filtered, demultiplexed single cell RNA sequencing samples.

Genotype	Conditions	Cell number	Median detected genes	Median counts per cells	Percent counts in mitochondrial genes
<i>Med12</i> wild type #1 (iG4)	2i + LIF	1082	5623	24292	5.09
<i>Med12</i> wild type #1 (iG4)	Dif	1450	6052	27890	4.20
<i>Med12</i> wild type #2 (iG6)	2i + LIF	1339	5663	24563	5.30
<i>Med12</i> wild type #2 (iG6)	Dif	1533	6024	28663	4.13
<i>Med12</i> mutant #1 (iG6)	2i + LIF	1362	5732	24834	5.43
<i>Med12</i> mutant #1 (iG6)	Dif	1763	5752	25342	4.43

Med12 mutant #2 (iG6)	2i + LIF	1257	5497	23361	5.52
Med12 mutant #2 (iG6)	Dif	1565	5805	25975	4.75
Med12 mutant #3 (iG6)	2i + LIF	1123	5528	23233	5.92
Med12 mutant #3 (iG6)	Dif	1396	5749	25697	4.81

Acknowledgments

I am deeply grateful for the chance to pursue my Ph.D. at MPI of Molecular Physiology in Dortmund. I sincerely thank all the people who contributed to the success of this project and made the journey an unforgettable and enjoyable experience.

First and foremost, I want to express my gratitude towards Dr. Christian Schröter. Your constant support, critical feedback, patience, and optimism have helped me tremendously during my journey. I appreciate the time you invested in discussing experimental ideas, results, and interpretations with me. Thank you for giving me the opportunity to be a part of your group.

I would like to extend my thanks to Prof. Dr. Jan Hengstler and Prof. Dr. Philippe Bastiaens for volunteering as my examiners. I am also grateful for all the feedback and discussions with Prof. Dr. Philippe Bastiaens and for the opportunity to be part of the Department for Systemic Cell Biology. Additionally, I received valuable advice during my thesis advisory committee meetings, and for that, I would like to thank Dr. Jochen Imig and Prof. Dr. Daniel Summerer.

I would also like to take this opportunity to express my gratitude towards two people who have had a significant impact on my decision to pursue a Ph.D. Thank you, Prof. Dr. Till Ischebeck, for providing me with the best start into my lab life that I could have possibly imagined. And thank you, Prof. Dr. Carsten Schultz, for giving me the incredible opportunity to experience and learn from two amazing scientific cultures at EMBL and OHSU.

I am grateful to Dr. Jochen Imig and his team, particularly Dr. Sama Shamloo and Stavroula Petroulia, for the preparation of the gRNA library and their assistance with the Bioanalyzer measurements. Your support saved me from weeks of trial and error.

Dr. Lucia Sironi and Christa Hornemann, you were fundamental for pursuing a Ph.D. at the MPI in Dortmund. I am very thankful for your valuable support.

At the MPI, I also received great service and materials for molecular biology procedures from the Protein Chemistry Facility. Within the Department for Systemic Cell Biology, I want to thank Dr. Peter Bieling for being such a fundamental part of many scientific discussions. Moreover, I want to thank all present and past members of the department, especially Dr. Sven Müller and Michael Schulz, and all technical assistances for making science as easy as possible.

I want to especially thank some of you: Starting with all members of the Schröterlab, Michelle, Julia, Sina, Marina, and Dhruv, you were a big factor in keeping me motivated for every day's lab adventures, thank you for your friendship. This expands to Yannic, Ankit, Lena, Mike, Hans, Farid and Svenja, thank you for every lunch break, even when discussing the same traffic situations over

and over again. I also want to thank all the temporary members of the Schröterlab, Pauliine, Fiorella, Caro, Pablo (finally, some basketball discussions in our office), and Kristin, who made this place even more exciting. Also, thank you, Jan-Erik, you were a big part of the welcome culture I experienced at the MPI. Finally, once more, I want to thank you, Michelle: Your help in the lab was extremely valuable.

Last but not least, I would like to thank my parents, my brother, and my partner for their constant support.