

RESEARCH

Open Access



Identification of X-chromosomal genes that drive sex differences in embryonic stem cells through a hierarchical CRISPR screening approach

Oriana Genolet¹, Anna A. Monaco^{1,2}, Ilona Dunkel¹, Michael Boettcher³ and Edda G. Schulz^{1*} 

* Correspondence: edda.schulz@molgen.mpg.de

¹Otto Warburg Laboratories, Max Planck Institute for Molecular Genetics, Berlin, Germany
Full list of author information is available at the end of the article

Abstract

Background: X-chromosomal genes contribute to sex differences, in particular during early development, when both X chromosomes are active in females. Double X-dosage shifts female pluripotent cells towards the naive stem cell state by increasing pluripotency factor expression, inhibiting the differentiation-promoting MAP kinase (MAPK) signaling pathway, and delaying differentiation.

Results: To identify the genetic basis of these sex differences, we use a two-step CRISPR screening approach to comprehensively identify X-linked genes that cause the female pluripotency phenotype in murine embryonic stem cells. A primary chromosome-wide CRISPR knockout screen and three secondary screens assaying for different aspects of the female pluripotency phenotype allow us to uncover multiple genes that act in concert and to disentangle their relative roles. Among them, we identify *Dusp9* and *Klh13* as two central players. While *Dusp9* mainly affects MAPK pathway intermediates, *Klh13* promotes pluripotency factor expression and delays differentiation, with both factors jointly repressing MAPK target gene expression.

Conclusions: Here, we elucidate the mechanisms that drive sex-induced differences in pluripotent cells and our approach serves as a blueprint to discover the genetic basis of the phenotypic consequences of other chromosomal effects.

Keywords: X chromosome, Gene dosage, Embryonic stem cells, Sex differences, Pluripotency, MAPK signaling, CRISPR screen, *Dusp9*, *Klh13*

Background

Chromosomal dosage can be altered through loss or gain of chromosomes, which, for autosomes, is generally associated with pathologies. Differential dosage of the mammalian sex chromosomes, by contrast, drives sex determination, in case of the Y, and contributes to sex differences, in case of the X chromosome [1, 2]. The dosage imbalance for X-chromosomal genes between XX females and XY males is largely neutralized in somatic cells through X-chromosome inactivation (XCI), where one X chromosome is



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

nearly completely silenced in each female cell [3]. A subset of genes escape XCI and likely contribute to sex differences, for example in the context of immunity and autoimmune diseases [4–6]. During early embryonic development, however, prior to the onset of XCI, the majority of X-linked genes are expressed at double the levels in female compared to male cells, resulting in substantial sex differences in cell state and developmental progression [7].

In many mammalian species, including mice, cows, and humans, female embryos develop more slowly than their male counterparts during early development [8]. Since no fetal hormones are produced at this stage, these observations have been attributed to variations in sex-chromosomal dosage, which in mice has been confirmed by the analysis of X-monosomic XO embryos [9, 10]. These sex differences have been investigated at the molecular level in female mouse embryonic stem cells (mESC), which are derived from early blastocyst embryos and thus carry two active X chromosomes. Female mESCs appear to be shifted towards a more naive ground state of pluripotency, which is associated with reduced activity of the differentiation-promoting MAP kinase (MAPK) signaling pathway, increased levels of (naive) pluripotency factors, and lower levels of global DNA methylation [11–13]. As a consequence, exit from the pluripotent state during differentiation is delayed in female compared to male mESCs [11]. Similar patterns have been observed in induced pluripotent stem cells (iPSCs) [12]. These X-dosage effects are likely mediated by X-encoded genes that modulate the stem cell state, the identity of which however remains mostly unknown. They might pose a biological checkpoint to ensure that only cells that have successfully inactivated one of their X chromosomes contribute to the differentiated adult organism.

In somatic cell types, MAPK signaling plays a key role in the regulation of cellular programs such as proliferation, but in mESCs, it drives the exit from the pluripotent state, while its inhibition stabilizes the self-renewing naive ground state of pluripotency [14, 15]. The main growth factors that stimulate MAPK signaling at these early developmental stages belong to the fibroblast growth factor (Fgf) family [16, 17]. Upon activation of the FGF receptor (FgfR), and the subsequent membrane recruitment of the growth factor receptor-bound protein 2 (Grb2), the small GTPase Ras is activated [18, 19]. Ras in turn triggers the kinase cascade of Raf, Mek, and Erk. Erk then translocates to the nucleus and activates MAPK target genes, including *Egr1* and *Spry4* (Additional file 1: Figure S1A) [20, 21]. Female mESCs express MAPK target genes at reduced levels compared to their male counterparts, suggesting an inhibition of the pathway [11]. To maintain homeostasis, the MAPK pathway is controlled by strong negative feedback loops on multiple levels [22]. MAPK inhibition therefore often leads to a counter-intuitive rise in phosphorylation levels of pathway intermediates due to reduced negative feedback activity [23, 24]. Female mESCs, where the MAPK pathway is inhibited, thus exhibit increased Mek phosphorylation compared to male cells, suggesting inhibition of the pathway downstream of Mek [11, 25].

MAPK signaling and pluripotency are tightly coupled, as the inhibition of this pathway blocks differentiation and leads to an increased expression of naive pluripotency markers and DNA hypomethylation, a hallmark of the naive pluripotent state [16, 26–28]. Reduced MAPK signaling in female mESCs thus results in increased levels of naive pluripotency factors such as *Nanog* and *Prdm14*, and global DNA hypomethylation [11, 13, 26, 29–34].

Although X-chromosomal dosage exhibits global effects on signaling and gene expression, central X-encoded genes that mediate these phenotypes remain to be uncovered [7]. The X-linked Erk phosphatase *Dusp9* has been shown to underlie sex differences in DNA methylation, since a heterozygous mutation resulted in DNA hypermethylation as observed in male cells [25, 35]. However, pluripotency factor expression and differentiation has been reported to be unaffected in such mutant cells [12]. Moreover, a series of other X-linked genes, including the transcription factors *Zic3* and *Tfe3* have been investigated, but their heterozygous deletion in female cells had no detectable effect [12]. Taken together, the genetic determinants that drive sex differences in mESCs remain incompletely understood.

We have performed a series of complementary CRISPR screens to identify X-linked genes that modulate MAPK signaling, pluripotency, and differentiation and found several genes that contribute to these phenotypes. We show that the E3 ubiquitin ligase adaptor protein *Klhl13* promotes pluripotency factor expression, while inhibiting MAPK target gene expression and differentiation. Female mESCs carrying heterozygous mutations of *Klhl13* and the known X-linked MAPK inhibitor *Dusp9* qualitatively recapitulate all aspects of the male pluripotency phenotype. We have thus identified the main drivers of X-dosage-dependent sex differences in mESCs and disentangled their relative contributions. Our approach can serve as a blueprint to investigate dosage effects of other chromosomes, such as those underlying trisomy 21, and our results will be important for development of gender-sensitive iPSC-based therapies.

Results

Pooled CRISPR knockout screen identifies X-chromosomal MAPK regulators

The X chromosome encodes ~1000 genes, any of which could potentially mediate the sex differences observed in murine pluripotent stem cells with respect to pluripotency factor expression, MAPK pathway activity, and differentiation efficiency [11, 12]. Since MAPK signaling represses pluripotency factors and promotes differentiation [16, 27, 28], we hypothesized that an X-linked MAPK inhibitor might underlie the female pluripotency phenotype [7, 11]. To comprehensively identify X-encoded MAPK inhibitors, we performed a chromosome-wide pooled CRISPR knockout screen (Fig. 1a). Through transduction of Cas9-expressing mESCs with an X-chromosomal sgRNA expression library, a pool of cells was generated with maximally one gene mutated per cell. Subsequent enrichment of cells with increased MAPK pathway activity and sequencing of their associated sgRNAs allowed identification of genes acting as MAPK inhibitors that, when deleted, increased MAPK signaling.

To be able to enrich cells with high MAPK activity through fluorescence-activated cell sorting (FACS), we generated a female mESC line (1.8-SRE-Elk), where expression of GFP was driven by a synthetic MAPK-sensitive SRE-Elk promoter (containing binding sites for the transcription factors *Elk1* and *Srf*, which are activated downstream of the MAPK pathway) (Fig. 1a). Reporter functionality was confirmed by treatment with an inhibitor of *Mek*, which resulted in the expected decrease in GFP fluorescence (Additional file 1: Figure S1B). To focus the screen on X-linked genes, we generated a custom sgRNA library (GeCKOx) containing a subset of sequences of the genome-wide GeCKO library [36], targeting 961 X-chromosomal genes with 6 sgRNAs per gene,

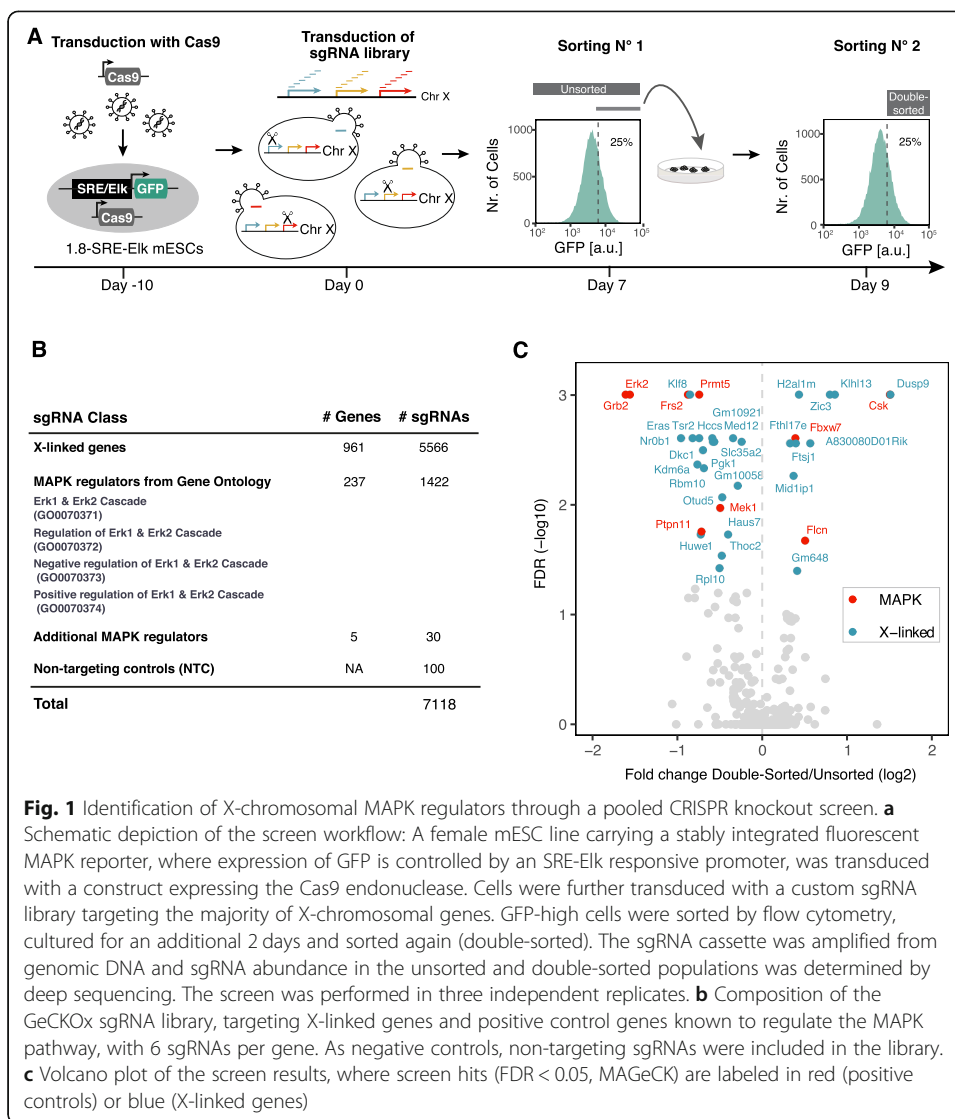


Fig. 1 Identification of X-chromosomal MAPK regulators through a pooled CRISPR knockout screen. **a** Schematic depiction of the screen workflow: A female mESC line carrying a stably integrated fluorescent MAPK reporter, where expression of GFP is controlled by an SRE-Elk responsive promoter, was transduced with a construct expressing the Cas9 endonuclease. Cells were further transduced with a custom sgRNA library targeting the majority of X-chromosomal genes. GFP-high cells were sorted by flow cytometry, cultured for an additional 2 days and sorted again (double-sorted). The sgRNA cassette was amplified from genomic DNA and sgRNA abundance in the unsorted and double-sorted populations was determined by deep sequencing. The screen was performed in three independent replicates. **b** Composition of the GeCKOx sgRNA library, targeting X-linked genes and positive control genes known to regulate the MAPK pathway, with 6 sgRNAs per gene. As negative controls, non-targeting sgRNAs were included in the library. **c** Volcano plot of the screen results, where screen hits (FDR < 0.05, MAGeCK) are labeled in red (positive controls) or blue (X-linked genes)

where possible (Additional file 2: Table S1). As controls, 237 genes implicated in MAPK pathway regulation according to gene ontology (GO) annotation and 100 non-targeting controls (NTC) were included in the library (Fig. 1b). Sequencing of the sgRNA library confirmed an even representation (Additional file 1: Figure S1C).

To investigate the female pluripotency phenotype, cells were generally grown in classical ESC culture conditions, containing Serum and LIF, if not stated otherwise. For the screen, 1.8-SRE-Elk mESCs were first transduced with a lentiviral vector expressing the Cas9 endonuclease, followed by blasticidin selection, sgRNA library transduction, and puromycin selection. After expansion for 7 days (5 days under selection), cells with high reporter activity were FACS-sorted, replated, and cultured for two additional days, before being sorted once again (Fig. 1a). We reasoned that such a double-sorting strategy would increase sensitivity of the screen. The sgRNA cassette was amplified from genomic DNA of all double-sorted (day 9) and unsorted (day 7) cell populations and sgRNA abundance in each sample was quantified by Illumina sequencing. SgRNA counts in all libraries were highly correlated and NTCs were neither enriched nor

depleted in the sorted fractions, suggesting that sufficient coverage was maintained at all steps of the screen (Additional file 1: Figure S1D-F).

Several core MAPK pathway components were significantly depleted in the GFP-high population (Erk2, Grb2, Frs2, Mek1, and Ptpn11), while *Csk*, a MAPK inhibitor [37], was enriched, showing that our screening setup could recover positive controls (Fig. 1c, red, Additional file 3: Table S2). Among the X-linked genes, 9 were significantly enriched and 18 were depleted in the sorted population (FDR < 0.05, MAGeCK, Fig. 1c, blue). *Dusp9*, *Klhl13*, and *Zic3* were the top-scoring MAPK inhibitors, and *Klf8*, *Nr0b1*, and *Eras* were the strongest activators (Fig. 1c).

In principle, enrichment in the double-sorted fraction at day 9 compared to the unsorted cells at day 7 could also be due to faster proliferation between the two sampling points. To identify genes that affect proliferation or viability, we compared sgRNA frequency in the cloned library and the unsorted cells at day 7 (Additional file 1: Figure S1G; Additional file 3: Table S2). Among the identified X-linked MAPK inhibitors, only *H2al1m* seemed to affect mESC proliferation positively, which would however lead to a decrease and not an enrichment, in sgRNA abundance between day 7 and day 9. In summary, we found a series of X-encoded inhibitors of the MAPK pathway, which might potentially drive the X-dosage-dependent pluripotency phenotype.

Secondary screens identify X-linked regulators of pluripotency factors, differentiation kinetics, and Mek phosphorylation

Having identified a set of putative X-linked MAPK pathway regulators, we further investigated their function in a series of complementary CRISPR screens. Specifically, we tested whether the identified candidate genes affected pluripotency factor expression, differentiation dynamics, and phosphorylation of Mek in a manner that would phenocopy the male pluripotency phenotype. For this purpose, a sub-library of the GeCKOx sgRNA library (GeCKOxs) was generated, targeting the 50 most enriched and depleted X-linked genes, together with the 10 most enriched and depleted MAPK controls from the primary screen (Fig. 2a; Additional file 1: Figure S2A; Additional file 4: Table S3). For each gene, the three most effective sgRNAs were selected. In addition, sgRNAs targeting 10 pluripotency regulators were included as further controls (*Sox2*, *Tbx3*, *Tcf3*, *Fgf2*, *Stat3*, *Esrrb*, *Tfcp2l1*, *Klf2*, *Nanog*, and *Oct4*).

To assess effects on pluripotency factor expression, we decided to assay for *Nanog* levels, which are consistently higher in female compared to male mESCs [11, 12, 25]. As a readout for differentiation efficiency, we monitored *Esrrb*, a naive pluripotency marker, which is downregulated with faster dynamics in cells with only one X chromosome [11]. We generated two transgenic mESC lines, where the endogenous *Nanog* and *Esrrb* genes, respectively, were tagged C-terminally with the fluorescent protein mCherry (Fig. 2b; Additional file 1: Figure S2B-C). Both reporters were downregulated upon differentiation, suggesting that they indeed mirrored expression of *Nanog* and *Esrrb*, respectively (Additional file 1: Figure S2D). In the pluripotency screen, we aimed at identifying *Nanog* activators, which when knocked out would reduce *Nanog* levels, and therefore sorted cells with low *Nanog* expression (Fig. 2c). Similarly, the differentiation screen aimed at identifying genes that would, when deleted, induce a more rapid downregulation of *Esrrb*. We therefore sorted *Esrrb*-low cells after 3 days of

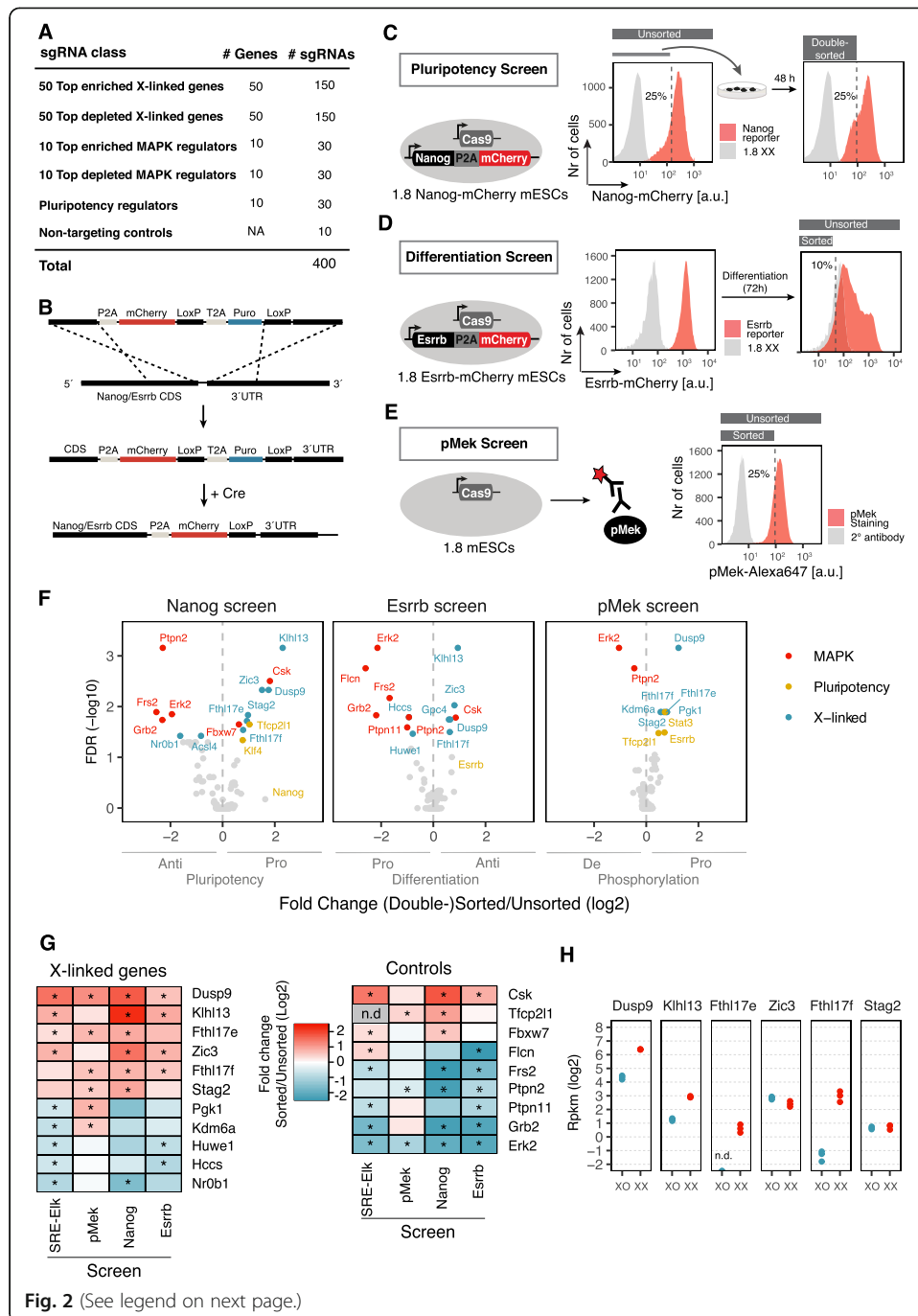


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Secondary CRISPR screens profiling pluripotency factor expression, differentiation kinetics and Mek phosphorylation. **a** Composition of the GeCKOxs sgRNA library, targeting hits from the primary MAPK screen and positive control genes with 3 sgRNAs per gene. **b** Schematic representation of the C-terminal tagging of the *Nanog* and *Esrrb* genes with the mCherry fluorescent protein through Cas9-mediated homologous recombination and subsequent Cre-mediated excision of the puromycin resistance cassette. *Nanog/Esrrb* and mCherry are linked through a P2A self-cleaving peptide. **c–e** Schematic depiction of the three secondary screens to profile effects on pluripotency factor expression (**c**), differentiation (**d**), and Mek phosphorylation (**e**). Female mESCs, carrying mCherry-tagged *Esrrb/Nanog* loci, as indicated, expressing the Cas9 endonuclease, were transduced with the sgRNA library in **a**. **c** In the Nanog screen, the 25% cells with the weakest mCherry fluorescence were enriched in two consecutive sorts (day 7 and day 9 after transduction). **d** For the *Esrrb* screen, cells were differentiated via LIF withdrawal for 3 days and the 10% cells with the lowest mCherry fluorescence were FACS sorted. **e** In the pMek screen, cells were stained intracellularly with a pMek-specific antibody and the 25% cells with the lowest signal were sorted. Three replicates were generated for the *Esrrb* and pMek reporter screens and two for the Nanog screen. **f** Volcano plots of the most enriched and depleted genes in the Nanog, *Esrrb*, and pMek screens. Genes with an FDR < 0.05 are highlighted as indicated. **g** Heatmap summarizing the results from all 4 screens. Enrichment of all X-linked (left) and control genes (right) that were significantly enriched or depleted in at least 2 screens is shown. *FDR < 0.05 (MAGeCK), n.d non-determined. **h** Expression levels for a subset of X-linked genes shown in **g** in 1.8XX and 1.8XO mESCs assessed by RNA sequencing

differentiation (Fig. 2d). For the Nanog screen, a double-sorting strategy similar to the primary MAPK screen was used, while only a single sorting step was performed for the differentiation screen, where a transient phenotype was analyzed (Fig. 2c, d).

In a third secondary screen, we aimed to test whether deletion of the candidate genes would result in decreased phosphorylation of Mek as observed in cells with one X chromosome [11, 25]. To this end, we performed an intracellular staining with a pMek-specific antibody and sorted cells with a low pMek signal (Fig. 2e). Staining specificity was confirmed by a higher pMek signal in XX compared to XO cells, together with an increase in pMek levels upon Mek treatment in the latter (Additional file 1: Figure S2E). Since the staining required cell fixation, only a single sorting step was possible. Sufficient sgRNA library representation was maintained throughout all steps of the screens (Additional file 1: Figure S2F). NTCs were neither enriched nor depleted in the pluripotency and differentiation screens, but seemed slightly but significantly depleted in the pMek screen (Additional file 1: Figure S2G). SgRNAs targeting the screen hits however exhibited a much stronger effect (Additional file 1: Figure S2G). Among the known MAPK regulators, the pathway components Erk2, Grb2, and Frs2 were identified as anti-pluripotency and pro-differentiation factors and the negative MAPK regulator Csk showed the opposite behavior (Fig. 2f, red, Additional file 3: Table S2). Erk2 also scored as the strongest negative regulator of Mek phosphorylation due to strong Erk-mediated negative feedback regulation [23, 24]. Ptpn2, a known negative regulator of MAPK signaling [38, 39], was surprisingly identified as an anti-pluripotency and pro-differentiation factor, potentially due to its previously reported inhibitory effect on Jak/STAT signaling, a pro-pluripotency pathway [40]. Moreover, Folliculin (Flcn) was identified as a strong pro-differentiation factor in agreement with its previously reported central role in early differentiation [41]. Finally, also the pluripotency factors Tfcp2l1 and Klf4 were identified as Nanog activators as expected [28, 42, 43]. *Nanog* and *Esrrb* themselves were enriched 3.1 (FDR = 0.67) and 1.6-fold (FDR = 0.2), respectively. The low statistical power to detect Nanog enrichment can be attributed to the fact that sgRNAs targeting *Nanog* become depleted over time, because they affect proliferation (Additional file 1: Figure S2H; Additional file 3: Table S2). Interestingly, the

pluripotency regulators Stat3, Esrrb, and Tfc2l1 scored as positive regulators of Mek phosphorylation, potentially in part due to crosstalk from the Jak/Stat to the MAPK signaling pathway [44]. In summary, all three secondary screens recovered known regulators supporting the validity of the approaches.

In all three secondary screens, 5–6 X-linked genes were enriched in the (double-)sorted populations, while only maximally 2 were depleted (Fig. 2f, blue, Fig. 2g, Additional file 3: Table S2). The only gene that significantly affected all 4 phenotypes (including the SRE-Elk screen, Fig. 1) was Dusp9, a known MAPK inhibitor that dephosphorylates Erk and has previously been implicated in sex differences in ES cells [12, 25, 35]. In addition, Klhl13, two members of the Fthl17 cluster, Fthl17e and Fthl17f, Zic3 and Stag2 significantly affected 2–3 phenotypes and generally showed the expected trend in all screens (Fig. 2g; Additional file 1: Figure S2I). Taken together, we have identified 6 genes that might contribute to the sex differences observed in mESCs, none of which, apart from Dusp9, has previously been implicated in mediating sex differences. Klhl13 encodes a substrate adaptor protein for the Cullin3 E3 ubiquitin-protein ligase complex with no known role in pluripotency or MAPK signaling regulation [45]. The Fthl17 gene cluster encodes ferritin-like proteins with unknown functions that are partially nuclear and lack ferroxidase activity [46]. Zic3 is a transcription factor implicated in pluripotency and early differentiation, whereas Stag2 regulates chromatin conformation and has also been shown to be involved in the maintenance of the pluripotent state in mESCs [47–50]. Among these candidates, the strongest effects were observed for Dusp9 and Klhl13.

Klhl13 and Dusp9 exhibit higher levels in females in vitro and in vivo

To further characterize the six identified putative mediators of the female pluripotency phenotype, we compared their expression pattern between cells with one and two X chromosomes, both in vitro and in vivo. We generated RNA sequencing data of the female mESC line used in all screens (1.8XX) and a subclone of that line with only one X chromosome (1.8XO, Additional file 5: Table S4). Although X-linked genes showed in general the expected 2-fold higher expression in XX compared to XO cells (Additional file 1: Figure S3A), two genes, Zic3 and Stag2, were expressed at similar levels in the two cell lines (mean fold-change 0.8 and 1.1 respectively), potentially due to gene-specific dosage-compensation mechanisms (Fig. 2h). Dusp9 and Klhl13 were expressed at 4.2- and 3.2-fold higher levels in XX compared to XO cells, respectively, and the two members of the Fthl17 cluster were essentially not expressed in the XO line (Fig. 2h). The strong expression difference for Fthl17e and Fthl17f can be explained by the fact that the cluster is maternally imprinted, such that it is only expressed from the paternal X chromosome, which is present only in female embryos and was probably also lost in the XO clone [51].

To assess expression patterns in mouse embryos in vivo, we analyzed epiblast cells in published single-cell RNA sequencing data collected between embryonic days E4.5 and E6.5 (Additional file 1: Figure S3B–G) [52]. Reactivation of the paternal X chromosome, which is silenced early in development in an imprinted form of XCI, is initiated around E4.5, completed at E5.5 and followed by random XCI around E6.5 [53, 54]. X-chromosomal expression was thus 1.6- and 1.4-fold higher in female compared to male

cells at E4.5 and E5.5, respectively, with the difference being largely neutralized by E6.5 (Additional file 1: Figure S3B). In contrast to mESCs, where both X chromosomes are active in the naive pluripotent state, *in vivo* naive pluripotency factors are primarily expressed prior to X reactivation around E3.5 and are mostly downregulated at E4.5 [53]. As a consequence, most naive markers were not well detected in the data set we analyzed and a combined analysis of 9 naive factors revealed only a slight trend towards higher expression in female cells at E4.5 (Additional file 1: Figure S3C-D). Analysis of a group of 9 markers of the primed pluripotent state, by contrast, showed a clear trend towards higher expression in all three time points (Additional file 1: Figure S3E-F), suggesting that differentiation of female cells with a double X-dosage is also delayed in embryos *in vivo*. Analysis of the six identified putative candidate genes revealed a trend towards higher expression in female cells at E5.5 for all factors, which was statistically significant ($p < 0.05$, Wilcoxon rank sum test) for *Dusp9*, *Fthl17e*, *Fthl17f*, and *Stag2* (Additional file 1: Figure S3G). In summary, all six factors were expressed at higher levels in female compared to male cells *in vivo*, but only four of them (*Dusp9*, *Klhl13*, *Fthl17e/f*) showed the same trend in the 1.8XX/XO cell lines *in vitro*. Since the 1.8XX/XO lines show a strong X-dosage-dependent phenotype [11], we concluded that the four differentially expressed factors would be the best candidates for mediating X-dosage effects on pluripotency and differentiation and decided to further validate *Dusp9* and *Klhl13*, which appeared to induce the strongest phenotypes.

Over-expression of *Klhl13* and *Dusp9* leads to an enhanced pluripotency state and slower differentiation kinetics in male mESCs

If *Dusp9* and *Klhl13* would indeed mediate the sex differences observed in mESCs, their over-expression in male cells should lead to a female-like pluripotency phenotype, while their heterozygous deletion should shift female cells towards a male-like phenotype. In order to over-express *Klhl13* and *Dusp9* from their endogenous loci in male mESCs, we implemented the CRISPR activation (CRISPRa) system. We made use of an E14 mESC line carrying the components of the CRISPRa SunTag system under control of a doxycycline-inducible promoter, which allows recruitment of multiple VP64 activation domains through a single sgRNA (Fig. 3a) [55, 56].

The SunTag system was recruited to either the *Dusp9* or *Klhl13* promoters using two different sgRNAs per gene and one sgRNA per cell line, leading to a 4- and 3.3-fold over-expression of *Dusp9* protein and to a 2.9- and 2.1-fold induction of *Klhl13* protein, respectively (Fig. 3b; Additional file 1: Figure S4A). We then characterized these cell lines with respect to pluripotency factor expression, differentiation dynamics, MAPK pathway activity, and global DNA methylation levels, all of which are affected by X-chromosomal dosage in mESCs. To assess MAPK pathway activity, we measured expression levels of *Spry4* and *Egr1* [20, 21], two well-known Erk target genes, by qPCR (Fig. 3c). Both MAPK target genes were strongly downregulated upon *Dusp9* over-expression (2.7/5.5-fold), as expected for an Erk phosphatase, while their expression was only slightly, albeit mostly not significantly reduced upon *Klhl13* over-expression. When assessing phosphorylation levels of MAPK pathway intermediates, we found that *Dusp9* over-expression reduced pErk levels 12-fold, again as expected for an Erk phosphatase, but increased pMek 22-fold, most possibly due to reduced negative feedback

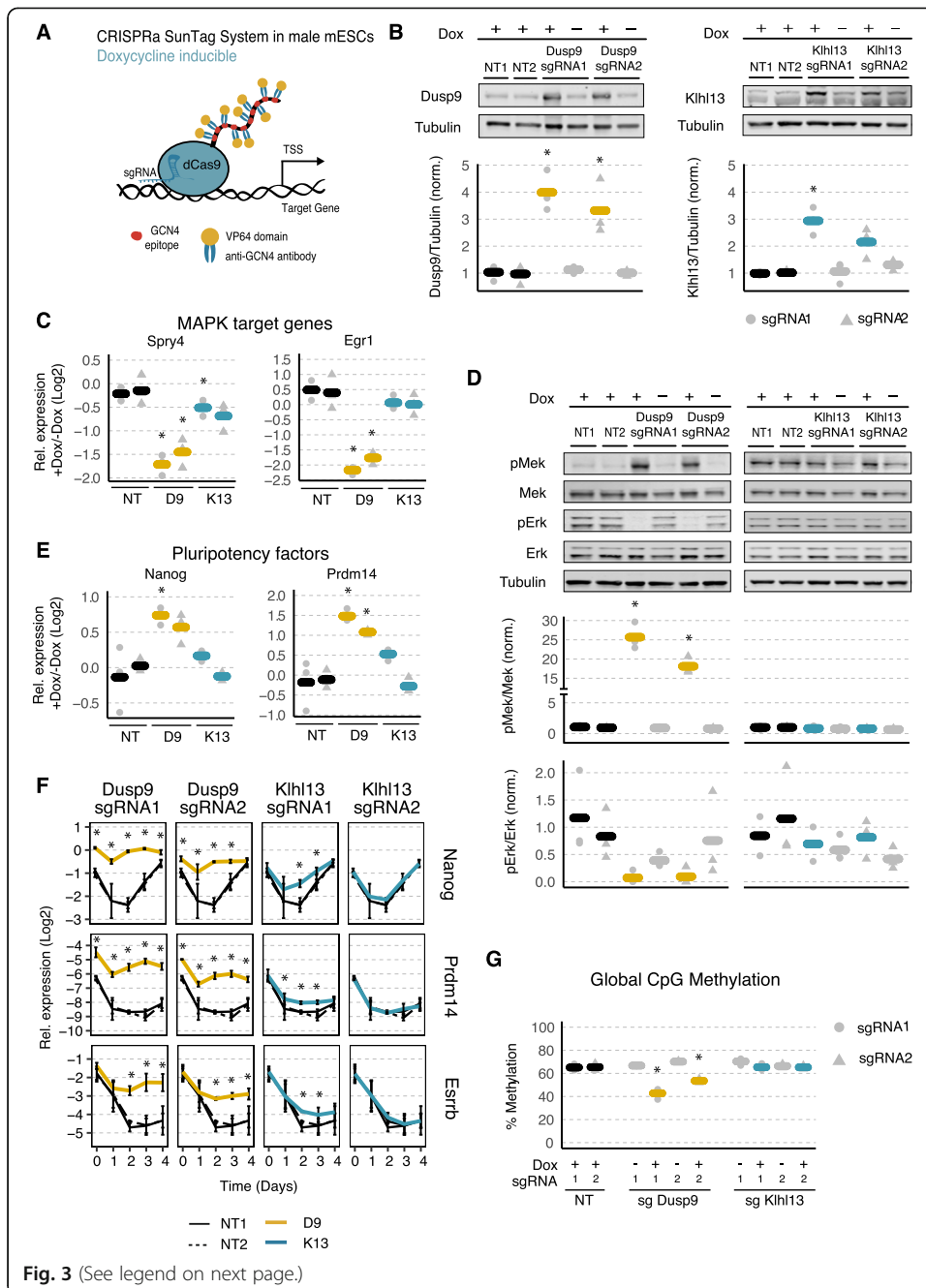


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Over-expression of *Klhl13* and *Dusp9* in male mESCs leads to an enhanced pluripotency state and slower differentiation kinetics. **a** Schematic representation of the dCas9-SunTag system used for gene activation. **b–e** To over-express *Dusp9* (yellow) and *Klhl13* (blue), male E14 mESCs, stably expressing the doxycycline-inducible SunTag system, were either transduced with one of two different sgRNAs targeting the respective promoter regions or with non-targeting control (NT) sgRNAs and were treated for 3 days with 1 $\mu\text{g/ml}$ doxycycline as indicated. Protein levels of *Dusp9* (left) and *Klhl13* (right) were quantified via immunoblotting (**b**), expression levels of MAPK target genes *Spry4* and *Egr1* (**c**) and of naive pluripotency factors *Nanog* and *Prdm14* (**e**) were assessed by qPCR and phosphorylation of Mek and Erk was quantified by immunoblotting (**d**). The immunoblot signals were normalized to Tubulin (**b**) or to total Mek/Erk (**d**) and to the mean of two doxycycline-treated non-targeting control sgRNAs. qPCR measurements were normalized to two housekeeping genes and to the respective untreated control (–Dox). Dots and triangles depict individual measurements of the two different sgRNAs, and thick bars show the mean of three biological replicates. **f** *Dusp9*- and *Klhl13* over-expressing mESCs were treated with 1 $\mu\text{g/ml}$ doxycycline 24 h before differentiation via LIF withdrawal for 4 days, and expression levels of pluripotency factors were measured by qPCR at different time points as indicated. Mean and standard deviation across 3 biological replicates is shown. **g** Global CpG methylation levels in cell lines over-expressing *Dusp9* and *Klhl13* via doxycycline treatment for 3 passages were assessed via pyrosequencing-based luminometric DNA methylation assay (LUMA). * $p < 0.05$ in a two-tailed paired Student's *t* test comparing the *Dusp9*/*Klhl13* over-expressing samples and the non-targeting controls (mean of sgRNA1 and sgRNA2)

inhibition (Fig. 3d). Over-expression of *Klhl13* by contrast had no significant effect on either Erk or Mek phosphorylation. A previous study had reported the opposite effect of *Dusp9* over-expression on Erk phosphorylation, potentially due to the requirement for trypsinization when analyzing feeder-dependent mESC lines, which were used in that study [25] (Additional file 1: Figure S4B). Taken together, these results confirm that *Dusp9* is a strong inhibitor of MAPK pathway activity, while *Klhl13* might slightly inhibit MAPK target gene expression, but does not affect pathway intermediates, which is in accordance with our screening results (Fig. 2g).

We next assessed how over-expression of *Klhl13* and *Dusp9* would affect pluripotency factor expression and differentiation dynamics. To this end, we quantified the pluripotency factors *Nanog* and *Prdm14*, which have been reported to be expressed at 2–4-fold higher levels in PSCs with two X chromosomes compared to those with one [11, 12, 25]. Over-expression of *Dusp9* in male mESCs resulted in a nearly comparable increase of *Nanog* (1.5–1.7-fold) and *Prdm14* (2.3–3-fold) levels. Upon *Klhl13* over-expression by contrast, only *Prdm14* was increased (1.6-fold) and only by the stronger sgRNA (Fig. 3e). A very similar trend was observed with regard to differentiation dynamics, where *Dusp9* over-expression essentially blocked downregulation of naive pluripotency markers, *Nanog*, *Prdm14*, and *Esrrb*, while for *Klhl13* only the stronger sgRNA had a mild effect on differentiation dynamics (Fig. 3f). Over-expression of *Dusp9* in male cells thus seemed to induce a strong shift towards the naive pluripotent state similar to female cells, while *Klhl13* over-expression resulted in only a minor shift.

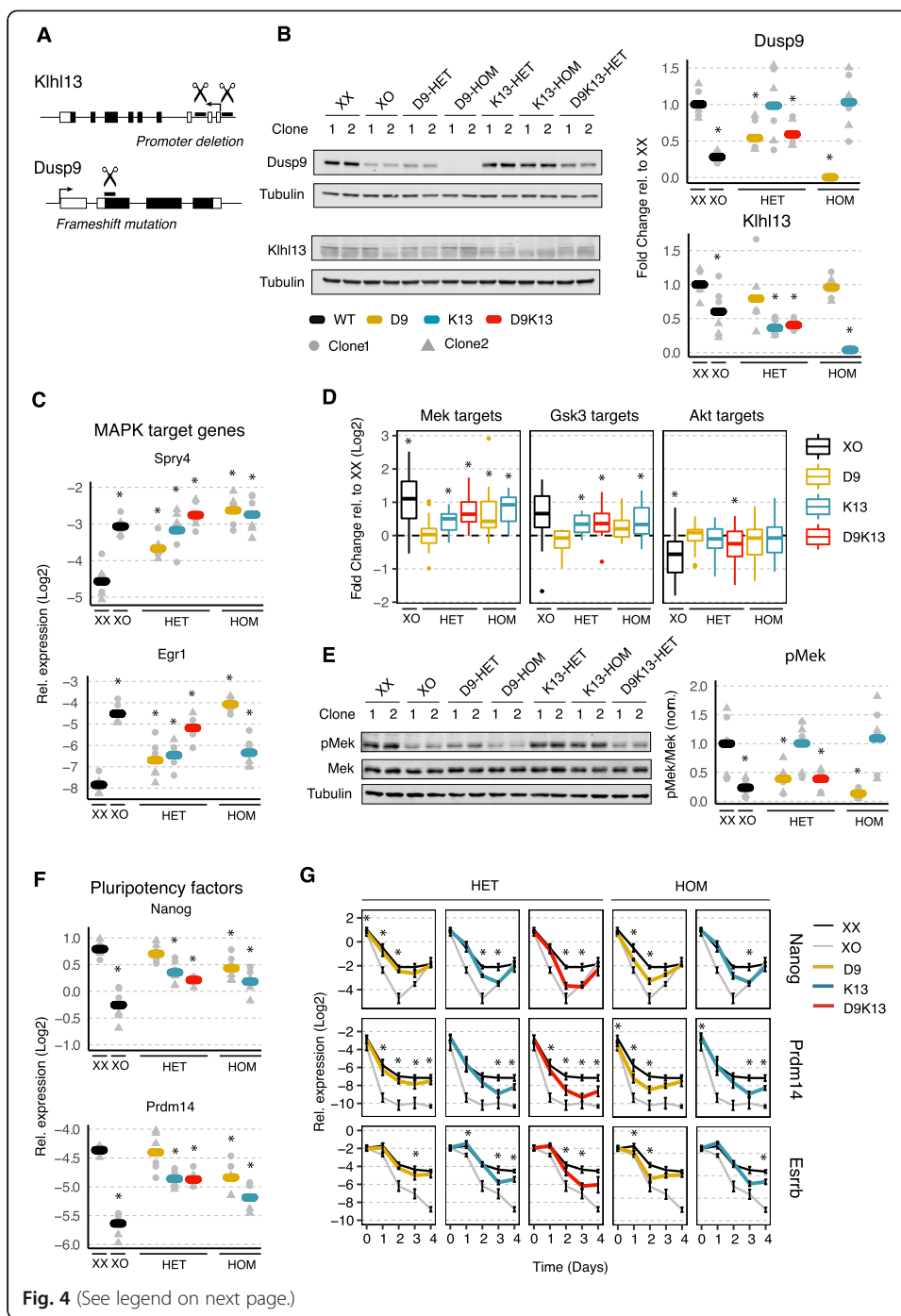
Since *Dusp9* has been suggested to be responsible for the reduction of global CpG methylation levels typically observed in female mESCs (20–30% compared to 60–80% in male mESCs) [13, 25, 31], we analyzed how over-expression of *Dusp9* and *Klhl13* affected global DNA methylation through the pyrosequencing-based luminometric DNA methylation assay (LUMA; Fig. 3g). Upon *Dusp9* over-expression, global DNA methylation levels were reduced from ~60% in NTC-transduced control cells to 53% and 42%, but were unaffected by *Klhl13* over-expression. Our results confirm a previously described effect of *Dusp9* on global DNA methylation [25].

Overall, we observe a stronger induction of a naive-like state in *Dusp9*- compared to *Klhl13*-over-expressing cells. It is important to note, however, that over-expression is less efficient for *Klhl13* than for *Dusp9* and that the observed effects seem to be strongly dose dependent. The fact that small, but significant effects are observed also for *Klhl13* with the stronger sgRNA (which increases *Klhl13* expression to levels similar albeit slightly lower compared to those in females) suggests that also *Klhl13* might contribute to sex differences with respect to pluripotency factor expression, differentiation, and MAPK target gene expression. To test this, we further investigated the role of both genes in female mESCs.

Mutation of one copy of *Klhl13* and *Dusp9* in female mESCs induces the male pluripotency state

If increased expression of *Klhl13* and *Dusp9* in female compared to male cells is indeed what drives sex differences in mESCs, their deletion on one X chromosome in female ESCs should induce the male phenotype. We therefore generated both heterozygous (HET) and homozygous (HOM) mutant mESC lines for *Klhl13* (K13) and *Dusp9* (D9) and a heterozygous double-mutant line (D9K13). For *Klhl13*, a 5-kb region spanning the promoter was deleted using Cas9, whereas for *Dusp9*, where attempts to create a promoter deletion were unsuccessful, frameshift mutations were introduced through an sgRNA targeting the start of the coding sequence (CDS) (Fig. 4a; Additional file 1: Figure S4C-D). Two clones were analyzed for each genotype throughout all experiments except for differentiation dynamics. Loss of *Klhl13* transcription in the respective mutants was confirmed by nascent RNA FISH (Additional file 1: Figure S4E) and all generated clones were karyotyped via double digest genotyping-by-sequencing (Additional file 1: Figure S4F) [57]. *Dusp9* protein levels were reduced ~1.8-fold in the respective HET mutants, which is less than the 3.5-fold reduction observed when comparing XX to XO cells, suggesting that *Dusp9* levels are modulated by other X-linked genes (Fig. 4b). In HET lines with a *Klhl13* mutation, the *Klhl13* protein was reduced ~2.7-fold (Fig. 4b). In all cell lines, we then analyzed MAPK signaling, pluripotency factor expression, differentiation, and global DNA methylation.

To assess whether MAPK pathway activity was affected in the mutant cell lines, we again quantified expression of the MAPK target genes *Egr1* and *Spry4* (Fig. 4c; Additional file 1: Figure S5A). Both were expressed at higher levels in all mutant lines compared to the XX control clones, suggesting that the MAPK pathway inhibition was at least partially lifted. Among the HET mutant lines, D9 showed the weakest effect, followed by K13 and D9K13, with the double mutant reaching similar expression levels as found in the XO control cells (3.4/6.5-fold for *Spry4*/*Egr1* in D9K13-HET vs 2.7/10.2 in XO). To get a more global picture of signaling activity, we analyzed a larger set of MAPK target genes using RNA sequencing (Fig. 4d; Additional file 1: Figure S5B; Additional file 6: Table S5). In agreement with the qPCR results, we found MAPK target genes were significantly increased in K13-HET cells and further elevated in D9K13 double mutants. We also assessed signatures of two other signaling pathways, Akt and Gsk3, implicated in pluripotency and differentiation, for which differential activity has been found in male and female mESCs [11, 58, 59]. Again, the heterozygous D9K13 mutant cells showed the strongest effects on Akt and Gsk3 target genes (Fig. 4d,



(See figure on previous page.)

Fig. 4 Heterozygous mutations of *Klhl13* and *Dusp9* in female mESCs partially phenocopy the male pluripotency state. **a** Schematic depiction of the strategies used to generate *Klhl13* (K13) and *Dusp9* (D9) mutant cell lines. **b–f** Comparison of female 1.8XX mESCs with a heterozygous (HET) or homozygous (HOM) deletion of *Dusp9* (yellow), *Klhl13* (blue) or both (red) with the parental XX line and XO controls (2 clones per genotype). Individual measurements are shown as gray dots (clone 1) and triangles (clone 2), and the mean across two clones and three biological replicates is indicated by a thick bar. **b** Immunoblot quantification of *Dusp9* (top) and *Klhl13* (bottom) protein levels, normalized to Tubulin and to the mean of the XX controls. **c** Quantification of MAPK target genes by qPCR. **d** Boxplots showing expression of *Mek* (left), *Gsk3* (middle), and *Akt* (right) target genes in cell lines with the indicated genotypes as assessed by RNA-seq. Boxes indicate the 25th to 75th percentiles, and the central line represents the median. **e** Quantification of pMek, normalized to total *Mek* and to the XX control cells by immunoblotting. **f** Pluripotency factor expression (*Nanog* and *Prdm14*) assessed by qPCR. **g** qPCR quantification of pluripotency factors during differentiation by 2i/LIF withdrawal in one clone for each genotype from the cell lines used in **b–f**. Mean and SD of three biological replicates is shown. * $p < 0.05$ Wilcoxon rank sum test (**d**), otherwise two-tailed paired Student's *t* test comparing each mutant/XO cell line and XX wildtype controls

Additional file 1: Figure S5B). It is important to note, however, that for none of the pathways we analyzed, target gene expression reached the levels found in XO control clones, suggesting that additional genes, other than *Dusp9* and *Klhl13*, are involved in their regulation.

We next investigated phosphorylation levels of *Mek* and observed a completely different pattern. Neither HOM nor HET mutations of *Klhl13* had any effect on pMek, but levels were reduced in the *Dusp9* mutants in a dose-dependent manner (Fig. 4e; Additional file 1: Figure S5C). The D9K13-HET double mutants resembled the D9-HET single mutants and exhibited a 2.6-fold pMek reduction compared to the wildtype XX control, thus approaching, but not reaching the 4.3-fold reduction observed in XO cells (Fig. 4e). Taken together, these results confirm the screening results that *Dusp9* and *Klhl13* both affect expression of MAPK target genes, but only *Dusp9* has a detectable effect on *Mek* phosphorylation (Fig. 2g), which is in accordance with their over-expression phenotypes in male mESCs (Fig. 3). These findings are in agreement with the role of *Dusp9* as an Erk phosphatase, which reduces Erk phosphorylation and consequently the Erk-mediated negative feedback upstream of *Mek*. Although *Dusp9* acts directly on the MAPK pathway, its deletion affects MAPK target gene expression less than the deletion of *Klhl13* (Fig. 4d), which is in contrast to results obtained from over-expression, where *Dusp9* shows stronger MAPK activation than *Klhl13* (Fig. 3c).

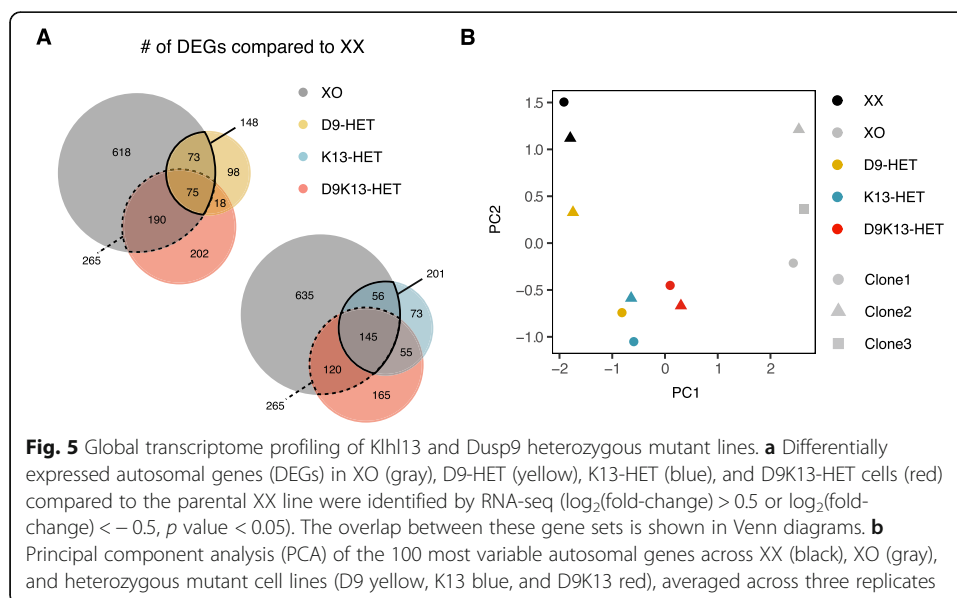
In the next step, we investigated pluripotency factor expression and differentiation kinetics. *Nanog* and *Prdm14* expression were significantly reduced in K13-HET, but not in D9-HET lines (Fig. 4f; Additional file 1: Figure S5D). D9K13 double mutant cells expressed similar levels as the K13 single mutant. With a ~1.5-fold reduction, the two genes could account for ~50% of the 2–2.4-fold decrease in *Nanog* and *Prdm14* levels observed in XO cells (Fig. 4f). For the assessment of differentiation dynamics, cells were first adapted to 2i conditions (with serum and LIF) for at least five passages. Since these conditions neutralize the expression differences of pluripotency factors between the cell lines in undifferentiated cells, they allow easier comparison of differentiation dynamics upon 2i/LIF withdrawal. Also here, *Klhl13* had a stronger effect than *Dusp9* (Fig. 4g). D9-HET mutants showed only a minimal reduction of *Esrrb*, *Nanog*, and *Prdm14* levels during differentiation compared to wildtype cells, while all three marker genes were reduced more in K13-HET cells (Fig. 4g). In the double D9K13 mutant, the effects of the

single mutants added up to nearly the levels observed in XO cells (Fig. 4g). We can conclude that *Klhl13* has a stronger effect on pluripotency factor expression and differentiation than *Dusp9* and that the double mutant can qualitatively, but not quantitatively reproduce the sex differences in mESC, suggesting that additional X-linked factors also contribute.

Finally, we also assessed global CpG methylation with the LUMA assay (Additional file 1: Figure S5E). In XX control cells, 31% of all CpG dinucleotides were methylated and levels were increased by ~10% in the single HET mutants and by ~15% in the HET double mutant and the HOM mutants. Given that D9K13-HET double mutants exhibited 44% methylation compared to 59% in XO control cells, *Klhl13* and *Dusp9* together could account for half of the differences seen in the XX/XO comparison.

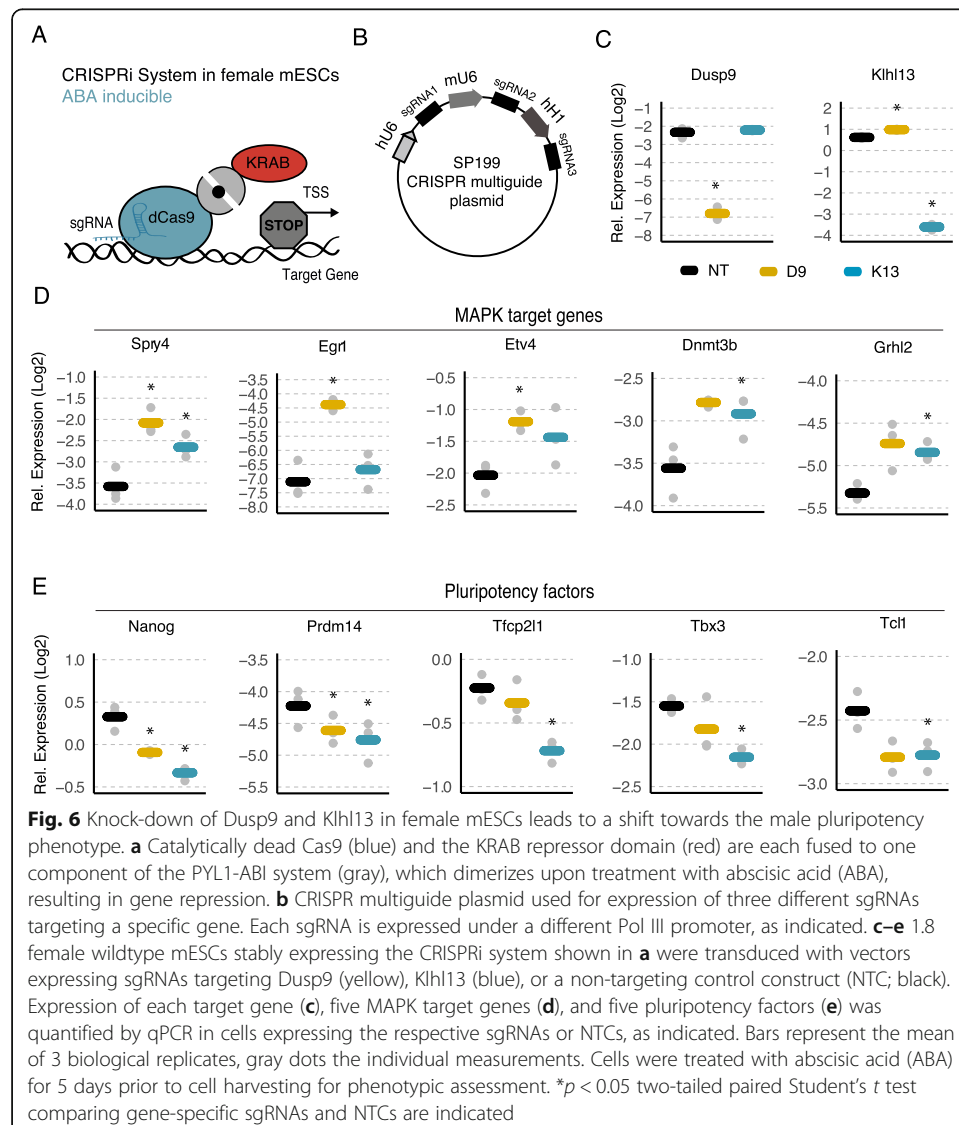
To get a more global picture of how well the mutant lines recapitulated the XO phenotype, we performed a transcriptome comparison. For each genotype, we identified autosomal genes that were differentially expressed when compared to the parental XX line. We found that 201 out of 956 differentially expressed genes (DEGs) on autosomes in XO cells were also differentially expressed in K13-HET cells, but only 148 in D9-HET lines (Fig. 5a). For the D9K13 double mutant, the overlap was 265 genes. A similar pattern was observed when performing principal component analysis (PCA). Also here the double mutant was found most closely to the XO controls, followed by K13-HET and D9-HET single mutants (Fig. 5b). These findings suggest that *Klhl13* contributes more to X-dosage induced transcriptome changes than *Dusp9* and that a combined effect of both can explain the observed sex differences best, but not completely.

When comparing the results of the mutant cell lines (Fig. 4) with the over-expression experiments in male cells (Fig. 3), it becomes apparent that the relative importance of the two genes seemed to be different in the two approaches. *Dusp9* had a much stronger effect than *Klhl13* on MAPK target genes and pluripotency factors in the over-expression experiment, while in the mutants, both genes affected MAPK target genes, but only *Klhl13* altered pluripotency factor expression. To distinguish, whether this discrepancy was due to the direction of the perturbation or different perturbations



strategies used, we implemented a third validation strategy, where *Dusp9* and *Klhl13* were downregulated through CRISPR interference (CRISPRi) in female mESCs, expressing an ABA-inducible split KRAB-dCas9 system (Fig. 6a). For both genes, 3 different sgRNAs targeting the gene's TSS were coexpressed from a single vector resulting in ~20-fold reduction of mRNA expression of each gene, compared to non-targeting control sgRNAs (Fig. 6b, c). Out of 5 quantified MAPK target genes, the majority was increased upon *Dusp9* and *Klhl13* repression, with somewhat stronger effects for *Dusp9* (Fig. 6d). The opposite pattern was observed, when profiling 5 naive pluripotency factors, where cells that downregulated *Klhl13* seemed to express consistently lower levels of these genes compared to cells with *Dusp9* downregulation (Fig. 6e), thus confirming the important role of *Klhl13* observed with the knockout approach (Fig. 4f).

In conclusion, multiple genes underlie the female pluripotency phenotype of which we have identified and validated a novel key player, *Klhl13*. *Dusp9* is responsible for the reduced levels of Mek phosphorylation in XX cells, but a combined effect of both genes together (partially) accounts for the global reduction of MAPK target genes in female



ES cells. The pluripotency and differentiation phenotypes by contrast can primarily be attributed to reduced *Klhl13* dosage in female cells. Since so far no mechanistic link between *Klhl13* and pluripotency or differentiation had been reported, we set out to investigate putative *Klhl13* interaction partners that might mediate the observed effects.

Identification of *Klhl13* interaction partners

Klhl13 is a member of the Cullin3 E3 ubiquitin ligase complex, where it acts as a substrate adaptor mediating protein ubiquitinylation, which might target proteins for proteasomal degradation (Fig. 7a) [60, 61]. We reasoned that the *Klhl13*-mediated sex differences we have identified might be due to reduced protein abundance of *Klhl13* substrate proteins in female compared to male cells, which affect pluripotency factors, differentiation, and MAPK target gene expression. To identify *Klhl13* substrates in mESCs, we profiled *Klhl13* interaction partners and then selected those with increased protein levels in K13-HOM mutant cells (Fig. 7b). To identify interaction partners, we ectopically expressed either full-length *Klhl13* or the substrate-binding Kelch domain, tagged with a green fluorescent protein (GFP), and identified binding partners by immunoprecipitation-mass spectrometry (IP-MS) using a GFP-specific antibody (Fig. 7b–d; Additional file 1: Figure S6A; Additional file 7: Table S6). Since E3 ubiquitin ligases usually interact with their substrates only transiently because they are rapidly degraded, the cells were treated with a proteasomal inhibitor for their stabilization. We identified a total of 197 interaction partners for the GFP-Kelch domain and 218 for full-length *Klhl13* that were enriched relative to the GFP-only controls (Fig. 7c, d; Additional file 7: Table S6). As expected, the interaction partners identified for full-length *Klhl13* and for the Kelch domain showed a large overlap, with 110 proteins being identified in both pull-downs. Two known interaction partners (*Nudcd3* and *Hsp90aa1*) were identified with both constructs and several members of the Cullin 3 complex (*Cul3*, *Klhl22*, *Klhl21*, *Klhl9*) were found to interact with full-length *Klhl13* only as expected (Fig. 7c, d, triangles) [62]. To identify putative *Klhl13* substrates among the 110 proteins found to interact with full-length *Klhl13* and with the Kelch domain only, we quantified the total proteome of K13-HOM cells and the parental XX control line through MS with label-free quantification (Fig. 7e; Additional file 8: Table S7). Among the 299 proteins that were significantly upregulated in the mutant cells (p value < 0.05, Benjamini-Hochberg FDR), 5 proteins (*Scml2*, *Peg10*, *Alg13*, *Larp1*, *Cct3*) had been identified as putative substrates in our IP-MS experiment (Fig. 7f).

We next investigated whether any of the five identified putative *Klhl13* target proteins would represent pro-differentiation factors by assessing MAPK target gene and pluripotency factor expression. To this end, we over-expressed them in female wildtype mESCs and tested whether their knock-down would rescue the phenotype of K13-HOM mutant cells. For the knock-down, we used the ABA-inducible split KRAB-dCas9 (CRISPRi) system described above (Fig. 6) and for gene over-expression an analogous system for gene activation, which recruits the VPR effector domain (CRISPRa) (Additional file 1: Figure S7A–B). Three different sgRNAs targeting the gene's TSS were coexpressed from a single vector. Perturbation strength, as assessed by qPCR, was

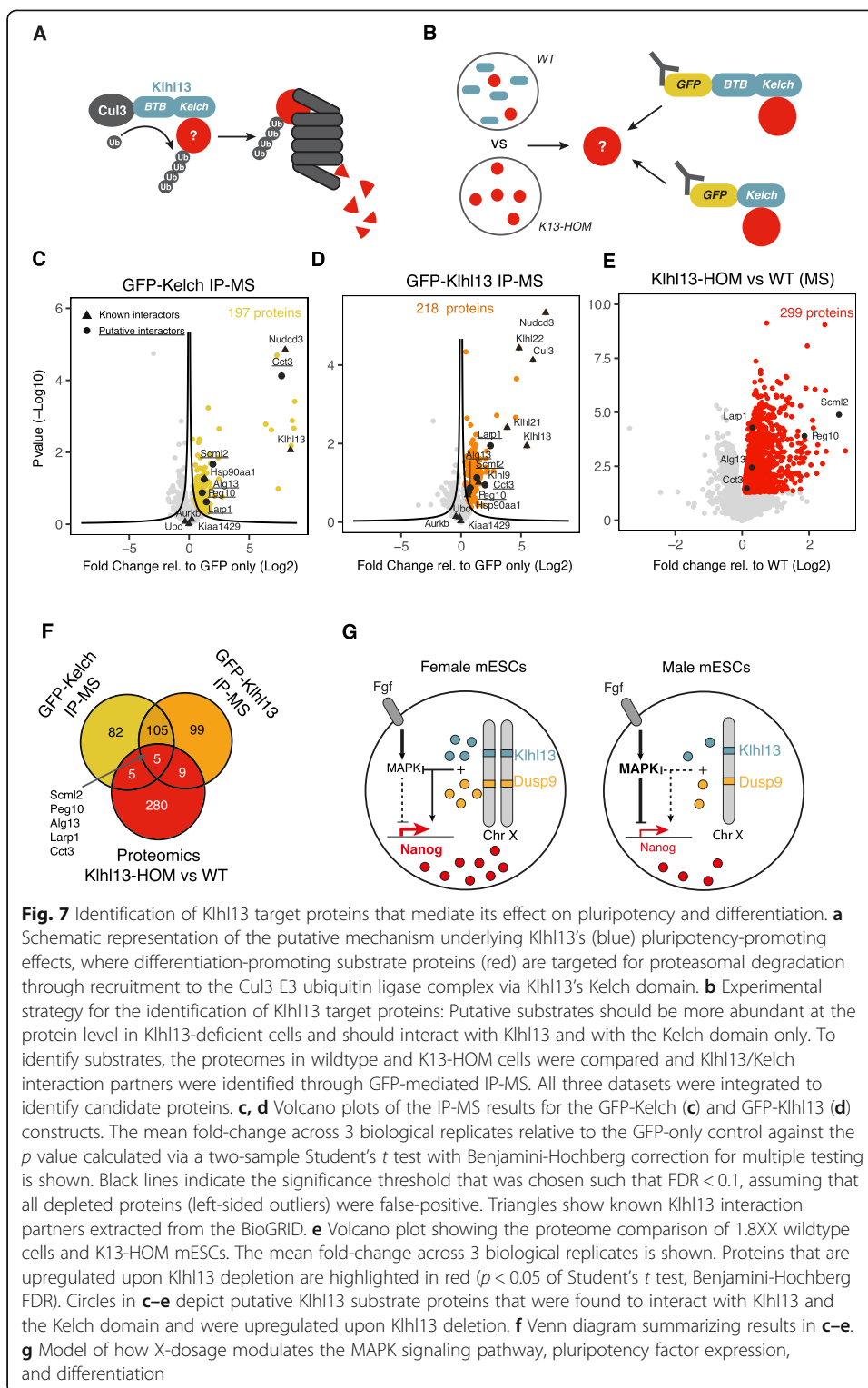


Fig. 7 Identification of Kih13 target proteins that mediate its effect on pluripotency and differentiation. **a** Schematic representation of the putative mechanism underlying Kih13's (blue) pluripotency-promoting effects, where differentiation-promoting substrate proteins (red) are targeted for proteasomal degradation through recruitment to the Cul3 E3 ubiquitin ligase complex via Kih13's Kelch domain. **b** Experimental strategy for the identification of Kih13 target proteins: Putative substrates should be more abundant at the protein level in Kih13-deficient cells and should interact with Kih13 and with the Kelch domain only. To identify substrates, the proteomes in wildtype and K13-HOM cells were compared and Kih13/Kelch interaction partners were identified through GFP-mediated IP-MS. All three datasets were integrated to identify candidate proteins. **c, d** Volcano plots of the IP-MS results for the GFP-Kelch (**c**) and GFP-Kih13 (**d**) constructs. The mean fold-change across 3 biological replicates relative to the GFP-only control against the p value calculated via a two-sample Student's t test with Benjamini-Hochberg correction for multiple testing is shown. Black lines indicate the significance threshold that was chosen such that $FDR < 0.1$, assuming that all depleted proteins (left-sided outliers) were false-positive. Triangles show known Kih13 interaction partners extracted from the BioGRID. **e** Volcano plot showing the proteome comparison of 1.8XX wildtype cells and K13-HOM mESCs. The mean fold-change across 3 biological replicates is shown. Proteins that are upregulated upon Kih13 depletion are highlighted in red ($p < 0.05$ of Student's t test, Benjamini-Hochberg FDR). Circles in **c-e** depict putative Kih13 substrate proteins that were found to interact with Kih13 and the Kelch domain and were upregulated upon Kih13 deletion. **f** Venn diagram summarizing results in **c-e**. **g** Model of how X-dosage modulates the MAPK signaling pathway, pluripotency factor expression, and differentiation

variable between genes, but reached at least 2-fold over-expression for all genes except *Larp1* in the CRISPRa experiment and a more than 2-fold reduction for all except *Cct3* upon CRISPRi (Additional file 1: Figure S7A-B).

We then assessed five MAPK targets (*Spry4*, *Egr1*, *Etv4*, *Dnmt3b*, and *Grhl2*) and five naive pluripotency markers (*Nanog*, *Prdm14*, *Tfcp2l1*, *Tbx3*, and *Tcl1*) by qPCR (Additional file 1: Figure S7C-D). For factors that mediate the *Klhl13* phenotype, we would expect an increase of MAPK target genes and a decrease in pluripotency markers upon over-expression, and the opposite trends upon knock-down. Generally, we only observed weak and mostly not consistent effects upon perturbation (Additional file 1: Figure S7C-D). However, a subset of factors exhibited some of the expected trends. *Peg10* led to a small, but significant increase in MAPK target genes and downregulation of naive pluripotency factors when over-expressed in female mESCs, a trend that was confirmed in an independent experiment with another sgRNA plasmid (Additional file 1: Figure S7E-G). This trend was however not observed in the CRISPRi experiment. Instead, knock-down of *Larp1* seemed to partially rescue the reduced pluripotency factor expression in K13-HOM cells. The reason why no effect was observed for *Larp1* in the CRISPRa experiment might be its inefficient over-expression in female wildtype mESCs (Additional file 1: Figure S7A).

In summary, we could not identify a single gene that might mediate the effects of *Klhl13* on the female pluripotency phenotype through the chosen approach. Instead, the phenotype might be mediated by several factors, potentially including *Peg10* and *Larp1*. Alternatively, ubiquitinylation of *Klhl13* substrates might lead to consequences other than proteasomal degradation, in which case also *Klhl13* interaction partners that were not upregulated in *Klhl13* knockout cells might be involved in *Klhl13* functions. We have thus narrowed down the list of candidate genes that warrant further investigation in the future.

Discussion

We present what, to our knowledge, is the first comprehensive functional identification of genes that drive phenotypic consequences of the loss or gain of an entire chromosome. We developed a hierarchical CRISPR screening approach, which allowed us to profile a large number of genes with respect to multiple phenotypes linked to sex differences in mESCs in an unbiased manner. In an initial X chromosome-wide screen, we identified a set of candidate genes, which were then further characterized for a role in modulating three additional molecular phenotypes. In this way, we identified several genes that potentially mediate X-chromosomal dosage effects and characterized the two strongest candidates in more detail, namely *Dusp9* and *Klhl13*. Through CRISPR-mediated over-expression in male and knockout or knock-down in female cells, we show that these two genes contribute to partially overlapping, yet distinct aspects of the X-dosage induced phenotype and that they appear to act in concert with additional factors. The X-dosage-dependent effects in pluripotent cells can thus not be attributed to a single X-linked gene, but arise from a complex interplay of multiple factors.

Dusp9 is a phosphatase that dephosphorylates the MAPK pathway intermediate Erk and is thus a known negative regulator of the pathway [35]. In agreement with previous reports, we found that *Dusp9* gain-of-function perturbations in male cells alter feedback strength and target gene expression [25, 63], while deletion of one copy of *Dusp9* in female cells results in the opposite phenotype. In addition, we confirmed the previously reported alterations in global DNA methylation levels, but the magnitude of the effects was less pronounced in our study, maybe due to differences in the cell lines,

culture conditions, and methylation assay used [25]. *Dusp9* also strongly affected pluripotency factor expression and differentiation in the gain-of-function experiments, which is consistent with a previous report [25]. We however observed only marginal effects in the female heterozygous *Dusp9* mutant cells, again in agreement with another study [12], which were however considerably stronger, when both copies of the gene were mutated or knocked-down.

For *Dusp9*, we thus observed a strong phenotype in the gain-of-function experiment in male cells and the opposite, albeit much weaker phenotype upon loss-of-function in female cells. Intriguingly, we found the opposite pattern for *Klhl13*, the second factor we investigated in detail. Here, the gain-of-function perturbation had only small effects, while loss-of-function led to an increase in MAPK target gene expression and a decrease for pluripotency factors, which was even more pronounced than the effects observed for *Dusp9*. This asymmetry between the gain- and loss-of-function perturbations remains puzzling and might point towards more complex interactions between multiple X-chromosomal factors.

Klhl13 is a substrate adaptor protein of the Cul3 E3 ubiquitin ligase complex [45] and has, to our knowledge, not yet been implicated in pluripotency, signaling, or X-dosage effects. Instead, it has been reported to be involved in mitotic progression through targeting Aurora kinase B in HeLa cells [45]. We could however generate mutant ES cells with a normal karyotype without difficulty, suggesting a different function for *Klhl13* in ES cells. While *Klhl13* did not affect phosphorylation of the MAPK pathway intermediate Mek, we found that knockout of only one copy of the gene resulted in a substantial increase in MAPK target gene expression, a reduction in pluripotency factors, and more efficient differentiation.

We hypothesized that a protein, which is targeted for proteasomal degradation through *Klhl13*-dependent ubiquitinylation might mediate the *Klhl13* phenotype. We therefore identified *Klhl13*-interacting proteins that were upregulated upon *Klhl13* deletion. While none of the five identified proteins could fully recapitulate the *Klhl13* phenotype, two of them, *Peg10* and *Larp1*, might contribute. *Peg10* is a known oncogene and has been shown to interact with Nanog and Oct4 in human cancer cells [64, 65], and *Larp1* is thought to regulate translation downstream of the mTOR pathway [66]. Identification of E3 ubiquitin ligase target proteins has also previously been reported to be challenging, probably due the transient nature of the interactions and the rapid target degradation [67]. Moreover, they typically have hundreds of substrates such that the *Klhl13* phenotype might be a combined effect of multiple target proteins. Another possibility is that ubiquitinylation might not lead to degradation, but might function as a signaling moiety instead [68, 69]. The *Klhl13* interaction partners we have identified that were not differentially expressed in *Klhl13* mutant cells might thus warrant further investigation.

Whatever the events downstream of *Klhl13* might be, or results clearly show that changes in Mek phosphorylation can be attributed completely to *Dusp9*, while *Klhl13* appears to be the main regulator underlying the X-dosage-dependent shift towards the naive pluripotent state. The combined effect of the two genes can thus account qualitatively for all aspects of the X-dosage-induced phenotype. The fact that the magnitude of effects in the double-mutant does not completely reproduce those seen in XO cells, suggests a contribution of one or several additional genes. Our screening approach has

identified some promising candidate genes that remain to be investigated in more detail. Moreover, additional screens in the D9K13 mutant background might allow identification of other factors in the future.

Among the other genes identified in the screens, *Zic3* is a pluripotency transcription factor that induces *Nanog* expression and enhances iPSC generation [47, 48, 70]. Accordingly, our screen identified *Zic3* as a MAPK inhibitor that promoted *Nanog* expression and impaired differentiation. Although it is not expressed in an X-dosage-sensitive manner in the cell line we used, higher expression in female compared to male cells has been reported in other cell lines [12]. However, pluripotency factor expression or differentiation was found to be unaffected in heterozygous *Zic3* mutant mESC lines [12]. Another factor identified in our screens is *Stag2*, a member of the Cohesin complex that has been implicated in the maintenance of pluripotency through mediating long-range regulation of pluripotency-associated genes [49]. A particularly interesting candidate is the *Fthl17* gene cluster, which contains seven genes that code for ferritin-like proteins, which however lack ferroxidase activity and are partially located in the nucleus [46]. They are maternally imprinted and therefore expressed only in female, but not in male blastocysts [51]. Their function is completely unknown, but the fact that they exhibit a female-specific expression pattern makes them intriguing candidates that warrant a more detailed investigation. Moreover, a contribution of X-linked imprinted genes to sex differences during development is also supported by the fact that mouse embryos with an XO genotype exhibit opposite growth phenotypes compared to XX embryos depending on whether they contain the maternal or paternal X chromosome [10].

Conclusions

In summary, we report central mechanisms underlying sex differences in murine pluripotent cells. The identified genes likely contribute to the X-dosage-dependent developmental delay in female embryos that has been reported in several mammalian species *in vivo* [9]. The X-dosage-induced stabilization of the naive pluripotent state might be important to ensure that X-dosage compensation has occurred before differentiation continues. Moreover, a specific differentiation speed might be required to ensure the establishment of exactly one inactive X chromosome in a female-specific manner [71]. After having identified the relevant genes, it will now be possible to investigate the functional relevance of X-dosage effects in pluripotent cells and developmental progression, both *in vitro* and *in vivo*.

Since the MAPK pathway plays an important role in cancer progression, our comprehensive profiling of X-encoded MAPK regulators might lead to a better understanding of the sex bias observed in certain cancer types [72]. Since loss of the inactive X chromosome or partial reactivation of X-linked genes in cancer has been observed in several studies [73, 74], reactivation of X-linked MAPK regulators might contribute to cancer susceptibility. In the context of gender medicine, sex differences in pluripotent cells are particularly relevant for therapeutic application of iPSCs. Although conventionally cultured human iPSCs retain an inactive X chromosome, it is often eroded with passage resulting in partial, but irreversible reactivation, which will be maintained during subsequent differentiation [75, 76]. Moreover, recently developed culture conditions for naive hiPS cells induce reversal of XCI [77, 78]. If a specific tissue is to be

regenerated through in vitro differentiation from iPSCs, we expect that double dosage of the genes identified in our study will modulate differentiation propensity in a sex-specific manner, which has already been described for the cardiac lineage [79]. Our results will thus enable a more focused investigation of how sex modulates iPSC therapy. In conclusion, our study is a first step towards understanding how X-dosage effects shape phenotypes in a sex-specific manner.

Methods

Molecular cloning

SRE-Elk reporter

For the generation of the pLenti-SRE/Elk-GFP-PEST-Hygro plasmid, a construct consisting of the MAPK-sensitive SRE-Elk promoter, containing repetitive binding sequences for the SRF (serum response factor), and Elk1 transcription factors (Sequence found in Additional file 9: Table S8), which drives the expression of a GFP protein fused to a destabilizing PEST sequence (kind gift from Morkel and Brummer lab), was cloned into the NheI and BsrGI (NEB) linearized hygromycin-resistant lentiviral vector lenti MS2-P65-HSF1_Hygro (Addgene 61426, [80]) by using the In-Fusion HD Cloning Kit (Takara Bio).

Repair template for the generation of Nanog and Esrrb reporters

Repair templates to tag Nanog and Esrrb with mCherry (pUC19-Nanog-mCherry-puro, pUC19-Esrrb-mCherry-puro, Additional file 9: Table S8) consisted of the P2A self-cleaving peptide followed by the mCherry coding sequence and a puromycin resistance cassette, flanked by ~ 400 bp homology regions to the Nanog/Esrrb locus (Esrrb-HA-Upstream: chr12:86,518,604-86,519,062, Esrrb-HA-Downstream: chr12:86,519,066-86,519,521, Nanog-HA-Upstream: chr6:122,713,142-122,713,552, Nanog-HA-Downstream: chr6:122,713,556-122,714,007, GRCm38/mm10 Assembly). All four fragments (upstream and downstream homology arms together with mCherry and the puromycin resistance cassette) were cloned into an XbaI (NEB) linearized pUC19 plasmid (Invitrogen) using the NEBuilder® HiFi DNA Assembly Cloning Kit (NEB) with 0.05 pmol/fragment and 10 µl of the NEBuilder master mix.

Klhl13 over-expression constructs

pLenti-PGK-Degron-GFP-Blast, pLenti-PGK-Degron-GFP-Klhl13-Blast, pLenti-PGK-GFP-Blast, and pLenti-PGK-GFP-Kelch, which were used to identify Klhl13 interaction partners, were generated and cloned into the pLenti-PGK-GFP-Blast lentiviral plasmid (Addgene 19,069, [81]) by GenScript (Additional file 9: Table S8). The Klhl13 isoform expressed in mESCs was used (ENSMUST00000115313.7). The Kelch domain (AA290 to AA585) was extracted from the SMART (<http://smart.embl-heidelberg.de/>) database. The Degron sequence consists of a mutated cytosolic prolyl isomerase FKBP12^{F36V}, but was not used in the reported experiments [82]. The GFP sequence was taken from previous publications [82].

sgRNA design

sgRNAs targeting the Nanog, Esrrb, Klhl13, and Dusp9 locus were designed using the CRISPR-Cas9 online tool <http://crispr.mit.edu:8079/>. Off-target scores based on in

silico quality and off-target predictions [83] were compared among the candidate sgRNAs and only the top-scoring ones were selected.

CRISPRa sgRNA sequences targeting the TSS of the *Dusp9*, *Cct3*, *Larp1*, *Peg10*, and *Scml2* genes were extracted from previously published libraries [84], whereas sgRNAs targeting the mESC-specific *Klhl13* isoform (ENSMUST00000115313.7) were designed using the CRISPR library designer (CLD) from the Boutros lab (Heigwer et al. 2016). sgRNAs targeting the TSS of the *Alg13* isoform ENSMUST00000238864.1 were designed using the CRISPOR sgRNA design tool [85].

CRISPRi sgRNA sequences targeting the TSS of the *Dusp9*, *Alg13*, *Cct3*, *Larp1*, *Peg10*, and *Scml2* were also extracted from previously published libraries [84, 86], whereas sgRNAs targeting the mESC-specific *Klhl13* isoform (ENSMUST00000115313.7) were designed using the CRISPOR sgRNA design tool [85]. Safe-targeting control sgRNAs were implemented for multiguide CRISPRa/CRISPRi validation experiments and sequences were extracted from previous publications [87].

sgRNA cloning

For sgRNA cloning into the PX330 (PX330-Nanog_sgRNA1, PX330-Nanog_sgRNA2, PX330-Esrrb_sgRNA1, PX330-Esrrb_sgRNA2, Additional file 9: Table S8, [88]) or PX458 plasmid (PX458-Dusp9_sgRNA1, [89]), two complementary oligos containing the guide sequence and a *BbsI* recognition site (Oligo F: 5' CACCGNNNNNNNNNN3' and Oligo R: 5' AAACNNNNNNNNNNC3') were annealed and cloned into the *BbsI* (NEB) digested target plasmid.

sgRNAs for CRISPRa (pU6-Dusp9.1-EF1Alpha-puro-T2A-BFP, pU6-Dusp9.2-EF1Alpha-puro-T2A-BFP, pU6-Klhl13.1-EF1Alpha-puro-T2A-BFP, pU6-Klhl13.2-EF1Alpha-puro-T2A-BFP, pU6-NTC.1-EF1Alpha-puro-T2A-BFP, pU6-NTC.2-EF1Alpha-puro-T2A-BFP) were cloned into a *BspI* and *BstXI* digested pU6-sgRNA-EF1a-puro-T2A-BFP plasmid (Addgene 60955, [90]) by annealing oligos containing the guide sequence and recognition sites for *BspI* and *BstXI* (Oligo F: 5' TTGGNNN...NNNGTTTAA-GAGC3' and Oligo R: 5' TTAGCTCTTAAACNNN...NNNCCAACAAG3') and ligating them together with the linearized vector using the T4 DNA ligase enzyme (NEB).

For the CRISPRi validation of *Dusp9* and *Klhl13* (SP199_multi_Dusp9_CRISPRi and SP199_multi_Klhl13_CRISPRi, Additional file 9: Table S8), as well as the CRISPRi/a validation of putative *Klhl13* interacting partners *Alg13*, *Cct3*, *Larp1*, *Peg10*, and *Scml2* (SP199_multi_NTC1, SP199_multi_NTC2, SP199_multi_Algl3_CRISPRi, SP199_multi_Cct3_CRISPRi, SP199_multi_Larp1_CRISPRi, SP199_multi_Peg10_CRISPRi, SP199_multi_Algl3_CRISPRa, SP199_multi_Cct3_CRISPRa, SP199_multi_Larp1_CRISPRa, SP199_multi_Peg10_CRISPRa, SP199_multi_Peg10_CRISPRa_2 and SP199_multi_Scml2_CRISPRa, Additional file 9: Table S8), three different sgRNAs targeting each gene were cloned into a single sgRNA expression plasmid with Golden Gate cloning, such that each sgRNA was controlled by a different Pol III promoter and fused to the optimized sgRNA constant region described in Chen et al. [91]. To this end, the sgRNA constant region of the lentiGuide-puro sgRNA expression plasmid (Addgene 52,963, [92]) was exchanged for the optimized version, thus generating the vector SP199 and the vector was digested with *BsmBI* (NEB) overnight at 37 °C and gel-purified. Two fragments were synthesized as gene blocks (IDT) containing the optimized sgRNA

constant region (handle) coupled to the mU6 or hH1 promoter sequences. These fragments were then amplified with primers that contained part of the sgRNA sequence and a BsmBI restriction site (primer sequences can be found in Additional file 9: Table S8) and purified using the gel and PCR purification kit (Macherey&Nagel). The vector (100 ng) and two fragments were ligated in an equimolar ratio in a Golden Gate reaction with T4 ligase and the BsmBI isoschizomer Esp3I for 20 cycles (5 min 37 °C, 20 min 20 °C) and transformed into NEB Stable competent *E. coli* [93]. Successful assembly was verified by Sanger sequencing.

Cell culture

Cell lines

Female 1.8 XX mESCs carry a homozygous insertion of 7xMS2 repeats in Xist exon 7 and are a gift from the Gribnau lab [11]. Several clones with XX or XO genotype (loss of one X chromosome) were generated through sub-cloning of the parental XX cell line. Female TX1072 ESCs carry a doxycycline responsive promoter in front of the *Xist* gene on one X chromosome and have been described previously [11]. For detailed information on the cell lines, refer to Additional file 9: Table S8. Low-passage Hek293T cells were a kind gift from the Yaspo lab.

The 1.8 SRE-Elk cell line was generated by lentiviral transduction of 1.8 XX mESCs with the pLenti-SRE/Elk-GFP-PEST-Hygro plasmid (Additional file 9: Table S8) followed by Hygromycin (250 ng/μl, VWR) selection. Single clones were picked and expanded and GFP expression confirmed via flow cytometry.

To identify Klhl13 interaction partners, female K13-HOM mESCs (Clone 2) were transduced with the lentiviral plasmids pLenti-PGK-Degron-GFP-Blast, pLenti-PGK-Degron-GFP-Klhl13-Blast, pLenti-PGK-GFP-Blast, and pLenti-PGK-GFP-Kelch plasmids (Additional file 9: Table S8) and selected using blasticidin (5 ng/μl, Roth). Protein expression was assessed via immunoblotting.

In 1.8-Nanog-mCherry and 1.8-Esrrb-mCherry reporter lines, the C-Terminus of the coding sequences of the Nanog or Esrrb genes, respectively, is tagged with the fluorescent protein mCherry, separated by a P2A self-cleaving peptide.

Cell lines over-expressing Klhl13 and Dusp9 via the CRISPRa Suntag system were generated by lentiviral transduction of E14-STN cells, which express the CRISPR activating Sun-Tag system [55] under a doxycycline-inducible promoter (kind gift from Navarro lab, [56]), with plasmids carrying sgRNAs targeted to the respective promoters or non-targeting controls (pU6-Klhl13.1-EF1Alpha-puro-T2A-BFP, pU6-Klhl13.2-EF1Alpha-puro-T2A-BFP, pU6-Dusp9.1-EF1Alpha-puro-T2A-BFP, pU6-Dusp9.2-EF1Alpha-puro-T2A-BFP, pU6-NTC.1-EF1Alpha-puro-T2A-BFP, pU6-NTC.2-EF1Alpha-puro-T2A-BFP, Additional file 9: Table S8; Additional file 1: Figure S3A) followed by puromycin selection (1 ng/μl, Sigma).

Dusp9 and Klhl13 heterozygous (HET) and homozygous (HOM) together with Dusp9 and Klhl13 double heterozygous mutant cell lines were generated via Cas9-mediated genome editing (see below) of 1.8 XX mESCs.

Cell lines for Klhl13 and Dusp9 knock-down were generated by lentiviral transduction of the 1.8 XX SP107 cell line (Clone A2, see below) with plasmids carrying sgRNAs targeting their respective promoters or a non-targeting control (SP199_multi_Dusp9_

CRISPRi, SP199_multi_Klh13_CRISPRi and SP199_multi_NTC1, Additional file 9: Table S8). Similarly, cell lines for Alg13, Cct3, Larp1, Peg10, and Scml2 knock-down and over-expression were generated by lentiviral transduction of the 1.8 XX K13-HOM SP107 and 1.8 XX SP106 cell line (see below), respectively, with plasmids carrying sgRNAs targeting their respective promoters (SP199_multi_Algl3_CRISPRi, SP199_multi_Cct3_CRISPRi, SP199_multi_Larp1_CRISPRi, SP199_multi_Peg10_CRISPRi, SP199_multi_Algl3_CRISPRa, SP199_multi_Cct3_CRISPRa, SP199_multi_Larp1_CRISPRa, SP199_multi_Peg10_CRISPRa, SP199_multi_Peg10_CRISPRa_2 and SP199_multi_Scml2_CRISPRa, SP199_multi_NTC1 and SP199_multi_NTC2, Additional file 9: Table S8). All cell lines were selected with puromycin (1 ng/ μ l, Sigma) for stable sgRNA integration.

Cell culture and differentiation

All mESC lines were grown without feeder cells on gelatin-coated flasks (Millipore, 0.1%) in serum-containing ES cell medium (DMEM (Sigma), 15% FBS (PanBiotech), 0.1 mM β -Mercaptoethanol (Sigma), 1000 U/ml leukemia inhibitory factor (LIF, Merck)). mESCs were passaged every second day at a density of 4×10^4 cells/cm² and medium was changed daily. Cells were differentiated by LIF withdrawal in DMEM supplemented with 10% FBS and 0.1 mM β -Mercaptoethanol at a density of 2×10^4 cells/cm² on fibronectin-coated dishes (Merck, 10 μ g/ml).

For the differentiation of mutant cell lines (Fig. 4g), cells were first adapted to 2i + LIF medium (ES cell medium with addition of 3 μ M Gsk3 inhibitor CT-99021 (Axon Medchem) and 1 μ M Mek inhibitor PD0325901 (Axon Medchem)) for at least five passages before undergoing differentiation via LIF withdrawal (see above). TX1072 XX and XO cells were grown in ES cell medium supplemented with 2i and differentiated by 2i/LIF withdrawal. Hek293T cells were cultured in DMEM supplemented with 10% FBS and passaged every 2 to 3 days.

Lentiviral transduction

For the generation of cell lines carrying randomly integrated transgenes using lentiviral transduction, DNA constructs were first packaged into lentiviral particles. For this, 1×10^6 Hek293T cells were seeded into one well of a 6-well plate and transfected the next day with the lentiviral packaging vectors: 1.2 μ g pLP1, 0.6 μ g pLP2, and 0.4 μ g VSVG (Thermo Fisher Scientific), together with 2 μ g of the desired construct using Lipofectamine 2000 (Thermo Fisher Scientific). Hek293T supernatant containing the viral particles was harvested after 48 h. 0.2×10^6 mESCs were seeded per 12-well and transduced the next day with 500 μ l of viral supernatant and 8 ng/ μ l polybrene (Sigma). Antibiotic selection was started 2 days after transduction and kept for at least 3 passages.

Genome editing

To generate 1.8-Nanog-mCherry and 1.8-Esrrb-mCherry reporter lines, 1×10^6 1.8 mESCs were transfected with 4 μ g of the pUC19-Nanog-mCherry-puro or pUC19-Esrrb-mCherry-puro plasmid (Additional file 9: Table S8) and 1.5 μ g of each of the sgRNAs plasmids (PX330-Nanog-sgRNA1/2 and PX330-Esrrb-sgRNA1/2) using 16.5 μ l of Lipofectamine 3000 and 22 μ l of P3000 (Thermo Fisher Scientific) according to the

manufacturer's recommendations. Cells were selected with puromycin (1 ng/ μ l, Sigma) for 3 days, starting at day 2 after transfection. The puromycin selection cassette was subsequently excised by transient transfection of a CRE recombinase expression plasmid pCAG-Cre (Addgene 13,775, [94]). Individual clones were expanded and tested for loss of puromycin resistance. mCherry fluorescence was measured via flow cytometry and clones were subsequently genotyped by PCR (Additional file 1: Figure S2B). All PCRs were carried out by using the Hotstart Taq Polymerase (Qiagen), a T_m of 56 °C and 30 cycles (Primer sequences are listed in Additional file 9: Table S8).

In order to generate *Klhl13* mutant mESCs, 4 guide RNAs were designed to target a 4.5-kb region around the *Klhl13* promoter (2 guide RNAs on each side) with the Alt-R[®] CRISPR-Cas9 System (IDT), which contains all necessary reagents for the delivery of Cas9-gRNA ribonucleoprotein complexes (RNP) into target cells. Briefly, crRNAs and tracrRNA (gRNA sequences in Additional file 9: Table S8) were mixed in equimolar concentrations and the 4 crRNAs and tracrRNA duplexes were subsequently pooled together. 2.1 μ l PBS, 1.2 μ l of the tra + cr duplex (100 μ M stock), 1.7 μ l Cas9 (61 μ M stock), and 1 μ l electroporation enhancer were pipetted together and incubated for 20 min. In total, 10⁵ cells were nucleofected with the mixture using the CP106 program of the Amaxa 4D-Nucleofector (Lonza) and plated on gelatin-coated 48-well plates. After 48 h, cells were seeded at a density of 10 cells/cm² into 10-cm plates. Individual clones were picked, expanded, and genotyped for the presence of the promoter deletion. The genotyping strategy is shown in Additional file 1: Figure S4B. For the amplification of the wildtype band, the HotStart Taq Polymerase (Qiagen) was used with an annealing temperature of 51 °C and 35 cycles. For the deletion, the Phusion HiFi Polymerase (NEB) was used with an annealing temperature of 63 °C and 35 cycles (Primer sequences are listed in Additional file 9: Table S8).

For the generation of *Dusp9* mutant mESCs, 2 \times 10⁶ WT and K13-HET (Clone 1) cells were nucleofected with 5 μ g of the PX458-*Dusp9*_sgRNA1 plasmid (Additional file 9: Table S8) and subsequently plated on gelatin-coated 6-cm plates. The next day, high GFP⁺ cells were single-cell sorted into a 96-well plate and expanded. Clones were screened for homozygous or heterozygous frameshift deletions via Sanger sequencing and immunoblotting. Heterozygous deletion of several selected clones was further confirmed via NGS. Briefly, a region surrounding the *Dusp9* deletion was amplified using the Phusion HiFi Polymerase (NEB) with a total of 30 cycles and an annealing temperature of 65 °C (Primer sequences in Additional file 9: Table S8, OG197/OG198). A second PCR using again the Phusion HiFi Polymerase (NEB) with a total of 14 cycles and an annealing temperature of 65 °C was performed in order to attach the Illumina adaptors and barcodes (Additional file 9: Table S8, OG202/OG210). A dual barcoding strategy was employed, where Illumina barcodes were included in the reverse and custom sample barcodes in the forward primers. Samples containing the same Illumina barcode but different custom sample barcodes were pooled in an equimolar fashion and sequenced on the Illumina Miseq platform PE150. Samples were aligned using Bowtie2 [95] and an index containing sample barcodes and possible deletion sequences based on previously generated Sanger sequencing data, gaining approximately 4000 reads per sample.

Generation of cell lines expressing the KRAB/VPR-dCas9 systems using piggybac transposition

The 1.8 XX SP107 (Clone A2) and 1.8 XX K13-HOM SP107 mESC lines stably express PYL1-KRAB-IRES-Blast and ABI-tagBFP-SpdCas9, constituting a two-component CRISPRi system, where dCas9 and the KRAB repressor domain are fused to ABI and PYL1 proteins, respectively, which dimerize upon treatment with abscisic acid (ABA) [96]. The 1.8 XX SP106 mESC line, on the other hand, expresses PYL1-VPR-IRES-Blast instead of PYL1-KRAB-IRES-Blast, together with ABI-tagBFP-SpdCas9, which leads to CRISPR-mediated activation of target genes when recruited to their TSS upon ABA treatment.

The 1.8 XX SP107, 1.8 XX K13-HOM SP107, and 1.8 XX SP106 mESC lines were generated through piggybac transposition. To this end, the puromycin resistance cassettes in the piggybac CRISPRi expression plasmid pSLQ2818 (pPB: CAG-PYL1-KRAB-IRES-Puro-WPRE-SV40PA PGK-ABI-tagBFP-SpdCas9, Addgene 84,241 [96]) and the CRISPRa expression plasmid pSLQ2817 (pPB: CAG-PYL1-VPR-IRES-Puro-WPRE-SV40PA PGK-ABI-tagBFP-SpdCas9, Addgene 84239 [96]) were exchanged for a blasticidin resistance, resulting in plasmid SP107 and SP106, respectively. The SP107 and SP106 plasmids were then, together with the hyperactive transposase (pBroad3_hyP-Base_IRES_tagRFP) [97], transfected into the 1.8 XX K13-HOM (Clone 1) and 1.8 XX (Clone 1) mESC lines, respectively, in a 1-to-5 transposase-to-target ratio. RFP-positive cells were sorted 24 h after transfection and cells were selected with blasticidin (5 ng/μl, Roth) for stable construct integration. After expansion, high BFP-positive cells were sorted. For the 1.8 XX SP107 mESCs, a clonal line was generated. Since target gene repression in cell lines stably expressing the SP107 construct transduced with sgRNAs was often observed already without ABA treatment, we could not make use of the inducibility of the system. Instead, 1.8 XX SP107 and 1.8 XX K13-HOM SP107 mESCs were always treated with ABA (100 μM) 5 days before the analysis and effects were compared to NTC sgRNAs. A 5-day ABA treatment (100 μM) was also carried out for the 1.8 XX SP106 mESC line prior to cell harvesting.

CRISPR KO screens

sgRNA library design

sgRNA sequences were extracted from the genome-wide GeCKO library [36]. For the GeCKOx library, a list of protein-coding and miRNA genes on the X chromosome was obtained from the NCBI Reference Sequence (Refseq) track on the UCSC genome browser [98, 99]. Control genes were included that were annotated with the Gene Ontology (GO) terms “Erk1 and Erk2 Cascade” (GO 0070371), “Regulation of Erk1 and Erk2 Cascade” (GO 0070372), “Negative regulation of Erk1 and Erk2 Cascade” (GO 0070373), and “Positive regulation of Erk1 and Erk2 Cascade” (GO 0070374). Additionally, known MAPK regulators Grb2, Fgfr2, Dusp5, Dusp7, and Dusp2 were added as additional controls. Six sgRNAs per gene and 100 non-targeting control sgRNAs were included in the GeCKOx library. The sgRNA sequences are provided in Additional file 2: Table S1. For the GeCKOx library, the 50 most enriched and depleted X-linked genes and the 10 most enriched and depleted MAPK regulators from the primary screen were identified using HitSelect [100]. The most enriched genes were identified

by comparing counts between Double-Sorted/Unsorted populations, whereas the most depleted genes were extracted by comparing counts between Unsorted/Double-Sorted populations. The 3 top-scoring sgRNAs for each gene were incorporated in the GeCKOxs library together with 10 non-targeting sgRNA controls. Additionally, 10 pluripotency regulators were added based on literature search (Sox2, Tbx3, Tcf3, Fgf2, Stat3, Esrrb, Tfcp2l1, Klf2, Nanog, Pou5f1). Klf4 was incorporated into the GeCKOxs library as a MAPK regulator (GO 0070373), having scored as a hit in the SRE/Elk screen, but was treated as a pluripotency factor control in later analyses. The sgRNA sequences are provided in Additional file 4: Table S3.

sgRNA library cloning

The GeCKOx and GeCKOxs sgRNA libraries were cloned into the lentiGuide-puro sgRNA expression plasmid (Addgene 52963, [92]). The vector was digested with BsmBI (NEB) overnight at 37 °C and gel-purified. sgRNA sequences were synthesized by CustomArray flanked with OligoL (TGGAAAGGACGAAACACCG) and OligoR (GTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGC) sequences. For the amplification of the library, 8 or 5 (GeCKOx/GeCKOxs) PCR reactions (Primer sequences in Additional file 9: Table S8, OG113/OG114) with approx. 5 ng of the synthesized oligo pool were carried out using the Phusion Hot Start Flex DNA Polymerase (NEB), with a total of 14 cycles and an annealing temperature of 63 °C in the first 3 cycles and 72 °C in the subsequent 11 cycles. The amplicons were subsequently gel-purified.

Amplified sgRNAs were ligated into the vector through Gibson assembly (NEB). Two 20 µl Gibson reactions were carried out using 7 ng of the gel-purified insert and 100 ng of the vector. The reactions were pooled, EtOH-precipitated to remove excess salts which might impair bacterial transformation, and resuspended in 12.5 µl H₂O. Nine microliters of the eluted DNA was transformed into 20 µl of electrocompetent cells (MegaX DH10B, Thermo Fisher Scientific) according to the manufacturer's protocol using the ECM 399 electroporator (BTX). After a short incubation period (1 h, 37 °C 250 rpm) in 1 ml SOC medium, 9 ml of LB medium with Ampicillin (0.1 mg/ml, Sigma) was added to the mixture and dilutions were plated in agar plates (1:100, 1:1000, and 1:10,000) to determine the coverage of the sgRNA libraries (600x for the GeCKOx and 2500x for the GeCKOxs). In total, 500 ml of LB media with Ampicillin was inoculated with the rest of the mixture and incubated overnight for subsequent plasmid purification using the NucleoBond Xtra Maxi Plus kit (Macherey-Nagel) following the manufacturer's instructions. To assess library composition by deep sequencing, a PCR reaction was carried out to add Illumina adaptors by using the Phusion High Fidelity DNA Polymerase (NEB), with an annealing temperature of 60 °C and 14 cycles (OG125/OG126). The PCR amplicon was gel-purified by using the Nucleospin Gel and PCR clean-up kit (Macherey-Nagel) following the manufacturer's instructions. Libraries were sequenced paired-end 50 bp on the HiSeq 2500 Platform yielding approximately 25 Mio fragments for the GeCKOx (20 pM loading concentration) and 1.3×10^6 fragments for the GeCKOxs library (22 pM loading concentration).

Viral packaging of sgRNA libraries

To generate virus carrying sgRNAs of the GeCKOx and GeCKOxs libraries, HEK293T cells were seeded into 12/8 (for GeCKOx/GeCKOxs) 10-cm plates and transfected the next day at 90% confluence. Each plate was transfected with 6.3 μg of pPL1, 3.1 μg of pLP2, and 2.1 μg of VSVG vectors (Thermo Fisher Scientific) together with 10.5 μg of the GeCKOx/GeCKOxs library plasmids in 1 ml of Opti-MEM (Life Technologies). Sixty microliters Lipofectamine 2000 Reagent (Thermo Fisher Scientific) were diluted in 1 ml Opti-MEM. Both mixtures were incubated separately for 5 min and then combined followed by a 20 min incubation, after which they were added dropwise to the HEK293T cells. Medium was changed 6 h after transfection. Transfected HEK293T cells were cultured for 48 h at 37 °C; afterwards, the medium was collected and centrifuged at 1800 \times g for 15 min at 4 °C. Viral supernatant was further concentrated 10-fold using the lenti-X™ Concentrator (Takara Bio) following the manufacturer's instructions and subsequently stored at -80 °C.

To assess the viral titer, 5 serial 10-fold dilutions of the viral stock were applied to each well of a 6-well mESC plate (MOCK plus 10⁻² to 10⁻⁶) for transduction with 8 ng/ μl polybrene (Merck). Two replicates were generated for each well. Selection with puromycin (1 ng/ μl , Sigma) was started 2 days after transduction and colonies were counted after 8 days. The number of colonies multiplied with the dilution factor yields the transducing units per ml (TU/ml), which ranged from 0.5–1.5 \times 10⁶ TU/ml.

Transduction and phenotypic enrichment

For the SRE-Elk screen, female 1.8-SRE-Elk mESCs were passaged twice before transduction with the lentiCas9 plasmid (Addgene 52,962, [92]). Blasticidin selection (5 ng/ μl , Roth) was started 2 days after transduction and kept for 4 passages, after which 6 \times 10⁶ cells were transduced with the sgRNA library (MOI = 0.3). Puromycin selection (1 ng/ μl , Sigma) was started 48 h after transduction and kept until harvesting at day 7 after transduction. The 25% of cells with the highest reporter activity were sorted. From these cells, 6–8 \times 10⁶ cells were snap-frozen and 6 \times 10⁶ were cultured for two additional days and subsequently sorted for GFP (top 25%). Around 8 \times 10⁶ unsorted cells were snap-frozen on day 7 and day 9 after transduction.

For the secondary screens, 2 \times 10⁶ female 1.8 XX Nanog-mCherry, 1.8 XX Esrrb-mCherry (see above for description), or 1.8 XX mESCs were transduced with the lentiCas9 plasmid as described above and then with the GeCKOxs library. 1.8 XX mESCs were stained for pMek (see below) on day 7 after transduction and the 25% of cells with the lowest pMek signal were sorted. 1.8-Esrrb-mCherry mESCs were differentiated by LIF withdrawal for 3 days starting on day 5 after transduction, after which cells were harvested and the 10% cells with the lowest mCherry fluorescence were sorted. 1.8-Nanog-mCherry mESCs were harvested on day 7 and the 25% cells with the lowest mCherry fluorescence were sorted. From these cells, around 2 \times 10⁶ were cultured for two additional days and subsequently sorted for mCherry (bottom 25%). Approximately 1 \times 10⁶ sorted and unsorted cells were snap-frozen for subsequent library preparation from all the secondary screens in order to maintain good library representation.

pMek intracellular staining

For the intracellular pMek staining, colonies were washed with PBS and dissociated to single cells with a 5 min trypsin (Life technologies) incubation. Trypsinization was stopped through addition of medium. Cells were disaggregated and pelleted, washed with PBS, and immediately fixed with 1.5% PFA (Roth). The cell mixture was incubated for 10 min at room temperature and subsequently centrifuged for 5 min at 500×g.

Cells were resuspended in ice-cold MeOH, incubated for 10 min on ice (0.5 ml/1 × 10⁶ cells) and centrifuged for 5 min at 500×g. Cells were washed once with staining buffer (PBS + 1% BSA (Sigma), 2 ml/1 × 10⁶ cells) and blocked for 10 min in staining buffer. Cells were incubated with the pMek-specific antibody (Cell Signaling, #2338, 1:100, antibodies are listed in Additional file 9: Table S8) for 30 min at room temperature (100 μl/1 × 10⁶ cells), then washed twice with staining buffer. Cells were then incubated with an anti-rabbit-Alexa647 antibody (Thermo Fisher Scientific, 1:400) for 15 min at room temperature (100 μl/1 × 10⁶ cells), washed twice with staining buffer before FACS sorting using the BD FACSAria™ II.

Preparation of sequencing libraries

For the SRE-Elk screen, genomic DNA was isolated from the frozen cell pellets using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's instructions. For the secondary screens, genomic DNA from frozen cell pellets was isolated via Phenol/Chloroform extraction due to higher yields. Briefly, cell pellets were thawed and resuspended in 250 μl of lysis buffer (1% SDS (Thermo Fisher Scientific), 0.2 M NaCl, and 5 mM DTT (Roth) in TE Buffer) and incubated overnight at 65 °C. In total, 200 μg of RNase A (Thermo Fisher Scientific) was added to the sample and incubated at 37 °C for 1 h. A total of 100 μg of Proteinase K (Sigma) was subsequently added followed by a 1 h incubation at 50 °C. Phenol/Chloroform/Isoamyl alcohol (Roth) was added to each sample in a 1:1 ratio, and the mixture was vortexed at RT for 1 min and subsequently centrifuged at 16000 × g for 10 min at room temperature. The aqueous phase was transferred to a new tube; 1 ml 100% EtOH, 90 μl 5 M NaCl, and 1 μl Pellet Paint (Merck) were added to each sample, mixed, and incubated at – 80 °C for 1 h. DNA was pelleted by centrifugation for 16,000×g for 15 min at 4 °C; pellets were washed twice with 70% EtOH, air-dried and resuspended in 50 μl H₂O.

The PCR amplification of the sgRNA cassette was performed in two PCR steps as described previously with minor modifications [36]. In order to ensure proper library coverage (300×), each sample was amplified in 6/2 PCR reactions (2 μg DNA/reaction) in the primary/secondary screens using the ReadyMix Kapa polymerase (Roche) with a total of 20 cycles and an annealing temperature of 55 °C (primer sequences in Additional file 9: Table S8, OG115/OG116).

Successful amplification was verified on a 1% agarose gel, and a second nested PCR was performed to attach sequencing adaptors and sample barcodes with 2.5 μl of the sample from the first PCR with a total of 11 cycles and an annealing temperature of 55 °C (OG125/OG126).

Resulting amplicons were loaded on a 1% agarose gel, purified using the Nucleospin Gel and PCR clean-up kit (Macherey-Nagel). Libraries from the primary screen were sequenced 2 × 50bp on the HiSeq 2500 Platform (18 pM loading concentration)

yielding 3–5 Mio fragments per sample. Secondary screens were sequenced 2×75 bp (Pluripotency and differentiation screens) on the Nextseq 500 (2.2 pM loading concentration) or 2×50 (pMek screen) on the HiSeq 2500 Platform (20 pM loading concentration) yielding at least 1.2×10^5 fragments per sample.

Data analysis

Data processing and statistical analysis was performed on the public Galaxy server usegalaxy.eu [101] with the MAGeCK CRISPR screen analysis tools [102, 103]. To this end, fastq files for read1 were uploaded to the Galaxy server. Alignment and read counting was performed with MAGeCK_count. Duplicated sgRNAs were excluded, leaving 6508 unique sgRNA sequences. Between 72 and 82% of reads mapped to the sgRNA library. Statistical analysis was performed with MAGeCK_test for each screen separately. Normalized counts and gene hit summary files were downloaded and analyzed in RStudio 3.5.3 using the stringr, tidyr, data.table, dplyr, and gplots packages. For easier interpretation of the results, common names were used instead of official gene symbols for a subset of genes (Erk2, Mek1, Fthl17e, Fthl17f, and H2al1m) in the figures and Additional file 3: Table S2. The 50 most enriched and depleted genes for the generation of the GeCKOxs sgRNA library from the primary screen were extracted using HitSelect [100].

DNA methylation profiling via LUMA

For the assessment of global CpG methylation levels, the luminometric methylation assay (LUMA) was performed as described previously [104]. For this, genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen) and 500 ng of DNA was digested either with HpaII/EcoRI (NEB) (Tube A) or MspI/EcoRI (NEB) (Tube B) in Tango Buffer (Thermo Fisher Scientific) in a total of 20 μ l for 4 h at 37 °C. Fifteen microliters of Pyrosequencing Annealing Buffer (Qiagen) was mixed with 15 μ l of each sample and overhangs were quantified by pyrosequencing using the following dispensation order GTGTGTCACACATGTGTGTG (nucleotides were pipetted in a two-fold dilution) in the PyroMark Q24 (Qiagen). The peak height from dispensation 13 (T) corresponds to the EcoRI digestion, and the peak height from dispensation 14(G) corresponds to the HpaII or the MspI digestion. For each sample, the HpaII/EcoRI ratio for tube A and the MspI/EcoRI ratio for tube B were calculated. The fraction of methylated DNA is then defined as: $1 - ((\text{HpaII/EcoRI}) / (\text{MspI/EcoRI}))$.

Flow cytometry

Cells were resuspended in Sorting buffer (1% FCS and 1 mM EDTA) or Staining Buffer (after pMek staining, PBS + 1% BSA) before flow cytometry, and cells were sorted using the BD FACSAria™ II. The sideward and forward scatter areas were used for live cell gating, whereas the height and width of the sideward and forward scatters were used for doublet discrimination. Analysis of FCS files was carried out using the FlowJo V10 Software (BD Biosciences). FCS files of the gated single-cell populations were visualized using RStudio and the Flowcore package.

Immunoblotting

Lysates were prepared from $\sim 2 \times 10^6$ cells by washing with ice-cold PBS and directly adding Bioplex Cell Lysis Buffer (Biorad) supplemented with the provided inhibitors. Plates were shaken at 4 °C at 300 rpm for 30 min, after which the lysates were transferred to 1.5 ml Eppendorf tubes and centrifuged at 4 °C and 4500×g for 20 min. Protein was transferred to a clean tube and quantified using the Pierce BCA kit (Thermo Fisher Scientific). For signaling proteins, 25 µg protein was applied per lane. For Dusp9, 10 µg and for Khl13 40 µg were loaded per lane. Proteins were transferred to nitrocellulose membranes by using the Trans-Blot Turbo Transfer System (Biorad) under semi-dry conditions. Membranes were blocked for 1 h with Odyssey Blocking Buffer/PBS (1:1) (Li-COR) at room temperature, followed by incubation with primary antibody (in Odyssey Blocking Buffer/PBST (1:1)) overnight at 4 °C. Signals were detected using near-infrared dye labeled secondary antibodies, and membranes were scanned using Li-COR Odyssey. Band intensities were quantified using the Image Studio Lite Ver 5.2 by calculating median intensities of the band area and subtracting the adjacent top/bottom background. Antibodies are listed in Additional file 9: Table S8.

RNA extraction, reverse transcription, qPCR

For gene expression profiling, $\sim 2 \times 10^6$ cells were washed with ice-cold PBS and lysed by directly adding 500 µl of Trizol (Invitrogen). RNA was isolated using the Direct-Zol RNA Miniprep Kit (Zymo Research) following the manufacturer's instructions. For quantitative PCR (qPCR), 1 µg RNA was reverse transcribed using Superscript III Reverse Transcriptase (Invitrogen) with random hexamer primers (Thermo Fisher Scientific) and expression levels were quantified in the QuantStudio™ 7 Flex Real-Time PCR machine (Thermo Fisher Scientific) using 2xSyBRGreen Master Mix (Applied Biosystems) normalizing to Rrm2 and Arpo. Primer sequences are listed in Additional file 9: Table S8.

RNA FISH

RNA FISH was performed as described previously with minor modifications [105]. Briefly, cells were singled out using Accutase (Invitrogen) and placed onto Poly-L-Lysine (Sigma)-coated (0.01% in H₂O, 10 min incubation at room temperature) coverslips #1.5 (1 mm) for 10 min. Cells were fixed in 3% paraformaldehyde in PBS for 10 min at room temperature and permeabilized for 5 min on ice in PBS containing 0.5% Triton X-100 and 2 mM Vanadyl-ribonucleoside complex (New England Biolabs). Coverslips were stored in -20 °C in 70% EtOH until further use.

Before incubation with the probe, the fixed cells were dehydrated through an ethanol series (80, 95, and 100% twice) and subsequently air-dried. BACs purified using the NucleoBond BAC kit (Macherey-Nagel) and spanning genomic regions of Huwel (RP24-157H12) and Khl13 (RP23-36505) were labeled by nick-translation (Abbot) using dUTP-Atto550 (Jena Bioscience) and Green dUTP (Enzo) respectively. Per coverslip, 60 ng probe was ethanol precipitated with Cot1 repeats (in order to suppress repetitive sequences in the BAC DNA that could hamper the visualization of specific signals), resuspended in formamide, denatured (10 min 75 °C), and competed for 1 h at 37 °C. Probes were co-hybridized in hybridization buffer overnight (50% formamide,

20% dextran sulfate, 2X SSC, 1 µg/µl BSA, 10 mM vanadyl-ribonucleoside). To reduce background, three 7 min washes were carried out at 42 °C in 50% formamide/2XSSC (pH 7.2) and three subsequent 5 min washes in 2X SSC at room temperature. Cells were stained with 0.2 mg/ml DAPI and mounted using Vectashield mounting medium for fluorescence (Vector Laboratories). Images were acquired using a widefield Z1 Observer (Zeiss) equipped with a × 100 objective and the filter set 38 and 43 (Zeiss). Image analysis was carried out using the Zen lite 2012 software (Zeiss).

Karyotyping

Cell lines were karyotyped via double digest genotyping-by-sequencing (ddGBS), a reduced representation genotyping method. The protocol was performed as described in the Palmers lab website, which was adapted from previously published protocols [57]. Briefly, the forward and reverse strands of a barcode adapter and common adapter were diluted and annealed, after which they were pipetted into each well of a 96-well PCR plate together with 1 µg of each sample and dried overnight (Oligo sequences are listed in Additional file 9: Table S8). The following day, the samples were digested with 20 µl of a NlaIII and PstI enzyme mix (NEB) in NEB Cutsmart Buffer at 37 °C for 2 h. After the digest, a 30 µl mix with 1.6 µl of T4 DNA ligase (NEB) was added to each well and placed on a thermocycler (16 °C 60 min followed by 80 °C 30 min for enzyme inactivation). By doing this, barcode and common adapters with ends complementary to those generated by the two restriction enzymes were ligated to the genomic DNA. Samples were cleaned with CleanNGS beads (CleanNA) using 90 µl of beads for each well and following the manufacturer's instructions. Samples were eluted in 25 µl ddH₂O and DNA was quantified using a dsDNA HS Qubit assay (Thermo Fisher). Samples were pooled in an equimolar fashion, size-selected (300-450 bp) by loading 400 ng of each pooled sample on an agarose gel followed by a cleaning step using the Nucleospin Gel and PCR clean-up kit (Macherey-Nagel). Samples were PCR amplified using the Phusion High-Fidelity DNA Polymerase (NEB) and an annealing temperature of 68 °C over 15 amplification cycles (OG218/OG219). Resulting amplicons were cleaned with CleanNGS beads in a 1:1.2 ratio (sample:beads) and sequenced with 2 × 75 bp on the MiSeq platform (12 pM loading concentration), yielding from 0.2 × 10⁶ to 1 × 10⁶ fragments per sample.

Data processing and statistical analysis was performed on the public Galaxy server usegalaxy.eu. Briefly, fastq files were uploaded and demultiplexed using the “Je-demultiplex” tool [106]. Reads were mapped to the mm10 mouse reference genome (GRCm38) using “Map with BWA” [107, 108]. Read counts for each chromosome were calculated with “multiBamSummary” [109] and normalized to a previously karyotyped XX control cell line (using Dapi-stained metaphase spreads and chromosome painting).

RNA-seq

For the RNA sequencing of 1.8 XX and 1.8 XO cell lines, libraries were generated using the Tru-Seq Stranded Total RNA library preparation kit (Illumina) with 1 µg starting material and amplified with 15 cycles of PCR. Libraries were sequenced 2 × 50 bp on one HiSeq 2500 lane (22 pM loading concentration), which generated ~ 40 Mio fragments per sample. The reads were mapped with the STAR aligner allowing for

maximally 2 mismatches to the mm10 mouse reference genome (GRCm38) and quantified using the ENSEMBL gene annotation [110], resulting in ~80% uniquely mapped reads. Gene expression values (rpkm) were obtained using the EdgeR package in RStudio [111]. Rpkms values are provided in Additional file 5: Table S4.

For RNA sequencing of the mutant cell lines, the QuantSeq 3' mRNA-Seq Library Prep Kit (FWD) for Illumina (Lexogen) was used with 800 ng starting material. Samples were sequenced with 1 × 75 bp on the NextSeq 500 Platform (2 pM loading concentration), which generated 1.5–3 Mio reads per sample. Read mapping, UMI collapsing, and read counting were performed with the FWD-UMI Mouse (GRCm38) Lexogen QuantSeq 2.6.1 pipeline from the BlueBee NGS data analysis platform (<https://www.bluebee.com/>). 62–68% of reads were mapped uniquely. Differential expression analysis was carried out, excluding X-chromosomal genes, using the EdgeR package in RStudio, together with normalization of gene expression values (cpm) [111]. Cpm values are provided in Additional file 6: Table S5.

Single-cell RNA-seq data analysis

For reanalysis of previously published scRNA-seq data, the normalized counts and the cell type annotation were downloaded from https://github.com/rargelaguet/scnmt_gastrulation. Sex annotation was provided by Ricard Argelaguet. For comparison of individual genes between male and female cells, a Wilcoxon rank sum test was performed using the `wilcox.test` function in R. For comparing chromosome-wide expression, counts for all genes located on a specific chromosome were summed up for each cell and then compared with a Wilcoxon rank sum test as described above. For the analysis of gene groups (naive and primed pluripotency markers), the log₂-transformed counts for all genes in the group were averaged for each cell and then analyzed as above.

Mass spectrometry

GFP immunoprecipitation

The GFP immunoprecipitation protocol was performed as described previously with minor modifications [112]. Briefly, cells were treated with 15 μM of MG132 for 3 h prior to harvesting. Cells were pelleted and resuspended in 1 ml of lysis buffer containing 150 mM NaCl, 50 mM Tris, pH 7.5, 5% glycerol, 1% IGEPAL-CA-630, 1 mM MgCl₂, 200 U benzonase (Merck), and EDTA-free complete protease inhibitor cocktail (Roche). Cells were incubated on ice for 30 min to allow cell lysis. Lysates were centrifuged at 4000×g and 4 °C for 15 min, and the supernatant was incubated with 50 μl magnetic beads coupled to monoclonal mouse anti-GFP antibody (Miltenyi Biotec) for 20 min on ice. Magnetic columns were equilibrated by washing first with 250 μl of 100% EtOH followed by two washes with the same volume of lysis buffer. After the 20 min incubation, the lysates were applied to the column followed by three washes with 800 μl of ice-cold wash buffer I (150 mM NaCl, 50 mM Tris, pH 7.5, 5% glycerol, and 0.05% IGEPAL-CA-630) and two washes with 500 μl of wash buffer II (150 mM NaCl, 50 mM Tris, pH 7.5, and 5% glycerol). Column-bound proteins were subsequently pre-digested with 25 μl 2 M urea in 50 mM Tris, pH 7.5, 1 mM DTT, and 150 ng trypsin (Roche) for 30 min at room temperature. Proteins were eluted by adding two times 50 μl elution buffer (2 M

urea in 50 mM Tris, pH 7.5, and 5 mM chloroacetamide). Proteins were further digested overnight at room temperature. The tryptic digest was stopped by adding formic acid to a final concentration of 2%.

Sample preparation for proteomics with label-free quantification (LFQ)

Proteomics sample preparation was done according to a published protocol with minor modifications [113]. Approximately 2×10^7 cells were lysed under denaturing conditions in a buffer containing 3 M guanidinium chloride (GdmCl), 5 mM tris (2-carboxyethyl) phosphine, 20 mM chloroacetamide, and 50 mM Tris-HCl pH 8.5. Lysates were denatured at 95 °C for 10 min shaking at 1000 rpm in a thermal shaker and sonicated in a water bath for 10 min. A small aliquot of cell lysate was used for the bicinchoninic acid (BCA) assay to quantify the protein concentration. In total, 50 µg protein of each lysate was diluted with a dilution buffer containing 10% acetonitrile and 25 mM Tris-HCl, pH 8.0, to reach a 1 M GdmCl concentration. Then, proteins were digested with LysC (Roche, Basel, Switzerland; enzyme to protein ratio 1:50, MS-grade) shaking at 700 rpm at 37 °C for 2 h. The digestion mixture was diluted again with the same dilution buffer to reach 0.5 M GdmCl, followed by a tryptic digestion (Roche, enzyme to protein ratio 1:50, MS-grade) and incubation at 37 °C overnight in a thermal shaker at 700 rpm.

LC-MS/MS instrument settings for shotgun proteome profiling

Peptide desalting was performed according to the manufacturer's instructions (Pierce C18 Tips, Thermo Scientific, Waltham, MA). Desalted peptides were reconstituted in 0.1% formic acid in water and further separated into four fractions by strong cation exchange chromatography (SCX, 3 M Purification, Meriden, CT). Eluates were first dried in a SpeedVac, then dissolved in 5% acetonitrile and 2% formic acid in water, briefly vortexed, and sonicated in a water bath for 30 s prior injection to nano-LC-MS. LC-MS/MS was carried out by nanoflow reverse phase liquid chromatography (Dionex Ultimate 3000, Thermo Scientific) coupled online to a Q-Exactive HF Orbitrap mass spectrometer (Thermo Scientific), as reported previously [114]. Briefly, the LC separation was performed using a PicoFrit analytical column (75 µm ID × 50 cm long, 15 µm Tip ID; New Objectives, Woburn, MA) in-house packed with 3-µm C18 resin (Reprosil-AQ Pur, Dr. Maisch, Ammerbuch, Germany). Peptides were eluted using a gradient from 3.8 to 38% solvent B in solvent A over 120 min at 266 nL per minute flow rate. Solvent A was 0.1% formic acid and solvent B was 79.9% acetonitrile, 20% H₂O, 0.1% formic acid. For the IP samples, a 1 h gradient was used. Nanoelectrospray was generated by applying 3.5 kV. A cycle of one full Fourier transformation scan mass spectrum (300–1750 m/z, resolution of 60,000 at m/z 200, automatic gain control (AGC) target 1×10^6) was followed by 12 data-dependent MS/MS scans (resolution of 30,000, AGC target 5×10^5) with a normalized collision energy of 25 eV. In order to avoid repeated sequencing of the same peptides, a dynamic exclusion window of 30 s was used. In addition, only peptide charge states between two to eight were sequenced.

Data analysis

Raw MS data were processed with MaxQuant software (v1.6.0.1) and searched against the mouse proteome database UniProtKB with 22,286 entries, released in December

2018. Parameters of MaxQuant database searching were a false discovery rate (FDR) of 0.01 for proteins and peptides, a minimum peptide length of seven amino acids, a first search mass tolerance for peptides of 20 ppm, and a main search tolerance of 4.5 ppm, and using the function “match between runs”. A maximum of two missed cleavages was allowed for the tryptic digest. Cysteine carbamidomethylation was set as fixed modification, while N-terminal acetylation and methionine oxidation were set as variable modifications. Contaminants, as well as proteins identified by site modification and proteins derived from the reversed part of the decoy database, were strictly excluded from further analysis.

Comparison of protein abundance for both the proteomics (K13-HOM vs XX wild-type) and the IP datasets (GFP-Kelch vs GFP, D-GFP-Klhl13 vs D-GFP) was performed with Perseus (v1.6.1.3). LFQ intensities, originating from at least two different peptides per protein group, were transformed by \log_2 . Only groups with valid values in at least one group were used, and missing values were replaced by values from the normal distribution. Statistical analysis for differential expression was done by a two-sample *t*-test with Benjamini-Hochberg (BH, FDR of 0.05) correction for multiple testing. The processed output files can be found in Additional file 7: Table S6 (IP-MS) and Additional file 8: Table S7 (Proteome).

For the identification of Klhl13 interaction partners, cut-offs were set from the data displayed in the volcano plots using a previously published method [115]. Briefly, a graphical formula as a smooth combination of the following parameters was implemented:

$$-\log_{10}(p) \geq \frac{c}{|x| - x_0}$$

x: enrichment factor of a protein

p: *p* value of the *t*-test, calculated from replicates

*x*₀: fixed minimum enrichment

c: curvature parameter

We optimized parameters *c* and *x*₀ such as to have 10% FDR (left-sided outliers) while maximizing the number of right-sided outliers. In the case of the GFP-Kelch IP, *c* = 0.32 and *x*₀ = 0.02. For the D-GFP-Klhl13 IP, *c* = 0.28 and *x*₀ = 0.04. Proteins without an associated gene name were filtered out in further analyses. Known Klhl13 interaction partners were extracted from the Biogrid database (Arih1, Aurkb, C1qbp, Cd2ap, Cops2, Cops4, Cops5, Cops6, Cops7a, Cul3, Dcun1d1, Hsp90aa1, Kiaa1429, Klhl21, Klhl22, Klhl9, Mad2l1, Nhlrc2, Nudcd3, Tfg, Ube2m, Ubxn7, Usp11, Zmym4).

Supplementary Information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13059-021-02321-2>.

Additional file 1: Figure S1. Identification of X-chromosomal MAPK regulators through a pooled CRISPR knockout screen. Figure S2. Secondary CRISPR screens profiling pluripotency factors, differentiation kinetics and Mek phosphorylation. Figure S3. Sex differences in embryonic stem cells and mouse embryos. Figure S4. Perturbation of Klhl13 and Dusp9 in mESCs. Figure S5. Heterozygous mutations of Klhl13 and Dusp9 in female mESCs partially phenocopy the male pluripotency state. Figure S6. Identification of Klhl13 target proteins that mediate its effect on pluripotency and differentiation. Figure S7. Effects of putative Klhl13 targets proteins on MAPK target gene and pluripotency factor expression.

Additional file 2: Table S1. GeCKOx sgRNA library.

Additional file 3: Table S2. Screen analysis: Enrichment, false discovery rate (FDR), rank and hit summary of the generated primary and secondary screens together with the proliferation effects for the SRE-Elk and Nanog screens.

Additional file 4: Table S3. GeCKOxs sgRNA library.

Additional file 5: Table S4. RNA-seq data of 1.8 XX/XO mESCs (Rpkms values).

Additional file 6: Table S5. RNA-seq data of Dusp9 and Kihl13 mutant mESCs (Cpm values).

Additional file 7: Table S6. IP-MS data for full length Kihl13 and the Kelch domain, showing LFQ (label-free quantification) protein intensities (log₂), fold change (FC) and statistical comparison of GFP-Kelch vs GFP and D-GFP-Kihl13 vs D-GFP.

Additional file 8: Table S7. Proteome comparison of 1.8 XX and K13HOM mESCs (2 clones, 3 replicates), including LFQ protein intensities (log₂), fold change and statistical comparison of K13HOM vs XX cells.

Additional file 9: Table S8. Antibodies, cell lines, plasmids, gRNAs, oligos and primers used in the study.

Additional file 10. Review history.

Acknowledgements

We would like to thank Nils Blüthgen, Heiner Schrewe, and Petra Knaus for the fruitful discussions; Joost Gribnau for providing the 1.8 XX cell line; Pablo Navarro for providing the E14 SunTag cell line; Jörn Schmiedel for sharing the sgRNA and repair plasmids for the generation of the Nanog and Esrrb reporter cell lines; Stephen Keyse for providing the anti-Dusp9 antibody; Luca Giorgetti for sharing the piggybac plasmid; Ricard Argelaguet for sharing the sex annotation of the scRNA-seq data set and the sequencing, mass spectrometry, and flow cytometry facility teams at the Max Planck Institute for molecular genetics, specially Dr. Bernd Timmermann, Dr. David Meierhofer, Sven Klages, Norbert Mages, Daniela Roth, Ilona Witzke, Myriam Hochradel, Beata Lukaszewska-McGreal, Dr. Claudia Giesecke-Thiel, and Uta Marchfelder.

Review history

The review history is available as Additional file 10.

Peer review information

Barbara Cheifet was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions

EGS and OG conceived the present work and planned the experiments. OG, AAM, and ID carried out the experiments. MB contributed to experimental planning and interpretation of results concerning the CRISPR knockout screens. EGS and OG took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript. The authors read and approved the final manuscript.

Funding

This work was supported by the Max Planck Research Group Leader program, E:bio Module III—Xnet grant (BMBF 031 L0072) and Human Frontiers Science Program (CDA-00064/2018) to E.G.S. O.G. is supported by the DFG (GRK1772, Computational Systems Biology). Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The datasets generated during the current study are available in the GEO repository, with identifiers GSE142348, GSE142349, and GSE142350 (SuperSeries GSE143784) [116] and via ProteomeXchange with identifiers PXD016729 [117] and PXD017875 [118]. The published single-cell RNA-seq data set [52] reanalyzed in this study is available at https://github.com/rargelaguet/scnmt_gastrulation [119].

Declarations

Ethics approval

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Otto Warburg Laboratories, Max Planck Institute for Molecular Genetics, Berlin, Germany. ²Present address: BIMSB, Max Delbrück Center for Molecular Medicine, Berlin, Germany. ³Medical Faculty, Martin-Luther-Universität Halle-Wittenberg, Halle, Germany.

Received: 9 February 2021 Accepted: 22 March 2021

Published online: 16 April 2021

References

1. Snell DM, Turner JMA. Sex chromosome effects on male-female differences in mammals. *Curr Biol*. 2018;28:R1313–24. <https://doi.org/10.1016/j.cub.2018.09.018>.

2. Ratnu VS, Emami MR, Bredy TW. Genetic and epigenetic factors underlying sex differences in the regulation of gene expression in the brain. *J Neurosci Res*. 2017;95:301–10. <https://doi.org/10.1002/jnr.23886>.
3. Schulz EG, Heard E. Role and control of X chromosome dosage in mammalian development. *Curr Opin Genet Dev*. 2013;23:109–15. <https://doi.org/10.1016/j.gde.2013.01.008>.
4. Schurz H, Salie M, Tromp G, Hoal EG, Kinnear CJ, Möller M. The X chromosome and sex-specific effects in infectious disease susceptibility. *Hum Genomics*. 2019;13:2. <https://doi.org/10.1186/s40246-018-0185-z>.
5. Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol*. 2016;16:626–38. <https://doi.org/10.1038/nri.2016.90>.
6. Wijchers PJ, Yandim C, Panosopoulou E, Ahmad M, Harker N, Saveliev A, et al. Sexual dimorphism in mammalian autosomal gene regulation is determined not only by Sry but by sex chromosome complement as well. *Dev Cell*. 2010;19:477–84. <https://doi.org/10.1016/j.devcel.2010.08.005>.
7. Schulz EG. X-chromosome dosage as a modulator of pluripotency, signalling and differentiation? *Philos Trans R Soc Lond B Biol Sci*. 2017;372 <https://doi.org/10.1098/rstb.2016.0366>.
8. Mittwoch U. Blastocysts prepare for the race to be male. *Hum Reprod*. 1993;8:1550–5. <https://doi.org/10.1093/oxfordjournals.humrep.a137889>.
9. Burgoyne PS, Thornhill AR, Boudrean SK, Darling SM, Bishop CE, Evans EP. The genetic basis of XX-XY differences present before gonadal sex differentiation in the mouse. *Philos Trans R Soc Lond B Biol Sci*. 1995;350:253–60 discussion 260. <https://doi.org/10.1098/rstb.1995.0159>.
10. Thornhill AR, Burgoyne PS. A paternally imprinted X chromosome retards the development of the early mouse embryo. *Development*. 1993;118:171–4.
11. Schulz EG, Meisig J, Nakamura T, Okamoto I, Sieber A, Picard C, et al. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell*. 2014;14:203–16. <https://doi.org/10.1016/j.stem.2013.11.022>.
12. Song J, Janiszewski A, De Geest N, Vanheer L, Talon I, El Bakkali M, et al. X-chromosome dosage modulates multiple molecular and cellular properties of mouse pluripotent stem cells independently of global DNA methylation levels. *Stem Cell Reports*. 2019;12:333–50. <https://doi.org/10.1016/j.stemcr.2018.12.004>.
13. Zvetkova I, Apedaille A, Ramsahoye B, Mermoud JE, Crompton LA, John R, et al. Global hypomethylation of the genome in XX embryonic stem cells. *Nat Genet*. 2005;37:1274–9. <https://doi.org/10.1038/ng1663>.
14. Lanner F, Rossant J. The role of FGF/Erk signaling in pluripotent cells. *Development*. 2010;137:3351–60. <https://doi.org/10.1242/dev.050146>.
15. Ying Q-L, Wray J, Nichols J, Battle-Morera L, Doble B, Woodgett J, et al. The ground state of embryonic stem cell self-renewal. *Nature*. 2008;453:519–23. <https://doi.org/10.1038/nature06968>.
16. Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*. 2007;134:2895–902. <https://doi.org/10.1242/dev.02880>.
17. Stavridis MP, Collins BJ, Storey KG. Retinoic acid orchestrates fibroblast growth factor signalling to drive embryonic stem cell differentiation. *Development*. 2010;137:881–90. <https://doi.org/10.1242/dev.043117>.
18. Kouhara H, Hadari YR, Spivak-Kroizman T, Schilling J, Bar-Sagi D, Lax I, et al. A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway. *Cell*. 1997;89:693–702. [https://doi.org/10.1016/s0092-8674\(00\)80252-4](https://doi.org/10.1016/s0092-8674(00)80252-4).
19. Lowenstein EJ, Daly RJ, Batzer AG, Li W, Margolis B, Lammers R, et al. The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell*. 1992;70:431–42. [https://doi.org/10.1016/0092-8674\(92\)90167-b](https://doi.org/10.1016/0092-8674(92)90167-b).
20. Hodge C, Liao J, Stofega M, Guan K, Carter-Su C, Schwartz J. Growth hormone stimulates phosphorylation and activation of elk-1 and expression of c-fos, egr-1, and junB through activation of extracellular signal-regulated kinases 1 and 2. *J Biol Chem*. 1998;273:31327–36. <https://doi.org/10.1074/jbc.273.47.31327>.
21. Casci T, Vinós J, Freeman M. Sprouty, an intracellular inhibitor of Ras signaling. *Cell*. 1999;96:655–65. [https://doi.org/10.1016/s0092-8674\(00\)80576-0](https://doi.org/10.1016/s0092-8674(00)80576-0).
22. Lake D, Corrêa SAL, Müller J. Negative feedback regulation of the ERK1/2 MAPK pathway. *Cell Mol Life Sci*. 2016;73:4397–413. <https://doi.org/10.1007/s00018-016-2297-8>.
23. Fritsche-Guenther R, Witzel F, Sieber A, Herr R, Schmidt N, Braun S, et al. Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. *Mol Syst Biol*. 2011;7:489. <https://doi.org/10.1038/msb.2011.27>.
24. Sturm OE, Orton R, Grindlay J, Birtwistle M, Vyshemirsky V, Gilbert D, et al. The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. *Sci Signal*. 2010;3:ra90. <https://doi.org/10.1126/scisignal.2001212>.
25. Choi J, Clement K, Huebner AJ, Webster J, Rose CM, Brumbaugh J, et al. DUSP9 modulates DNA hypomethylation in female mouse pluripotent stem cells. *Cell Stem Cell*. 2017;20:706–719.e7. <https://doi.org/10.1016/j.stem.2017.03.002>.
26. Ficiz G, Hore TA, Santos F, Lee HJ, Dean W, Arand J, et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell*. 2013;13:351–9. <https://doi.org/10.1016/j.stem.2013.06.004>.
27. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, et al. Nanog is the gateway to the pluripotent ground state. *Cell*. 2009;138:722–37. <https://doi.org/10.1016/j.cell.2009.07.039>.
28. Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*. 2009;460:118–22. <https://doi.org/10.1038/nature08113>.
29. Ooi SK, Wolf D, Hartung O, Agarwal S, Daley GQ, Goff SP, et al. Dynamic instability of genomic methylation patterns in pluripotent stem cells. *Epigenetics Chromatin*. 2010;3:17. <https://doi.org/10.1186/1756-8935-3-17>.
30. Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T, Grabole N, et al. Naive pluripotency is associated with global DNA hypomethylation. *Nat Struct Mol Biol*. 2013;20:311–6. <https://doi.org/10.1038/nsmb.2510>.
31. Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HHD, Matarese F, et al. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell*. 2013;13:360–9. <https://doi.org/10.1016/j.stem.2013.06.002>.
32. Hackett JA, Dietmann S, Murakami K, Down TA, Leitch HG, Surani MA. Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Reports*. 2013;1:518–31. <https://doi.org/10.1016/j.stemcr.2013.11.010>.

33. Milagre I, Stubbs TM, King MR, Spindel J, Santos F, Krueger F, et al. Gender differences in global but not targeted demethylation in iPSC reprogramming. *Cell Rep.* 2017;18:1079–89. <https://doi.org/10.1016/j.celrep.2017.01.008>.
34. Pasque V, Karnik R, Chronis C, Petrella P, Langerman J, Bonora G, et al. X chromosome dosage influences DNA methylation dynamics during reprogramming to mouse iPSCs. *Stem Cell Reports.* 2018;10:1537–50. <https://doi.org/10.1016/j.stemcr.2018.03.019>.
35. Caunt CJ, Keyse SM. Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling. *FEBS J.* 2013;280:489–504. <https://doi.org/10.1111/j.1742-4658.2012.08716.x>.
36. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* 2014;343:84–7. <https://doi.org/10.1126/science.1247005>.
37. Okada M. Regulation of the SRC family kinases by Csk. *Int J Biol Sci.* 2012;8:1385–97. <https://doi.org/10.7150/ijbs.5141>.
38. van Vliet C, Bukczynska PE, Puryer MA, Sadek CM, Shields BJ, Tremblay ML, et al. Selective regulation of tumor necrosis factor-induced Erk signaling by Src family kinases and the T cell protein tyrosine phosphatase. *Nat Immunol.* 2005;6:253–60. <https://doi.org/10.1038/ni1169>.
39. Mattila E, Pellinen T, Nevo J, Vuoriluoto K, Arjonen A, Ivaska J. Negative regulation of EGFR signalling through integrin- α 1 β 1-mediated activation of protein tyrosine phosphatase TCPTP. *Nat Cell Biol.* 2005;7:78–85. <https://doi.org/10.1038/ncb1209>.
40. Yamamoto T, Sekine Y, Kashima K, Kubota A, Sato N, Aoki N, et al. The nuclear isoform of protein-tyrosine phosphatase TC-PTP regulates interleukin-6-mediated signaling pathway through STAT3 dephosphorylation. *Biochem Biophys Res Commun.* 2002;297:811–7. [https://doi.org/10.1016/S0006-291X\(02\)02291-x](https://doi.org/10.1016/S0006-291X(02)02291-x).
41. Betschinger J, Nichols J, Dietmann S, Corrin PD, Paddison PJ, Smith A. Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell.* 2013;153:335–47. <https://doi.org/10.1016/j.cell.2013.03.012>.
42. Martello G, Bertone P, Smith A. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *EMBO J.* 2013;32:2561–74. <https://doi.org/10.1038/emboj.2013.177>.
43. Hall J, Guo G, Wray J, Eyres I, Nichols J, Grotewold L, et al. Oct4 and LIF/Stat3 additively induce Krüppel factors to sustain embryonic stem cell self-renewal. *Cell Stem Cell.* 2009;5:597–609. <https://doi.org/10.1016/j.stem.2009.11.003>.
44. Cacalano NA, Sanden D, Johnston JA. Tyrosine-phosphorylated SOCS-3 inhibits STAT activation but binds to p120 RasGAP and activates Ras. *Nat Cell Biol.* 2001;3:460–5. <https://doi.org/10.1038/35074525>.
45. Sumara I, Quadroni M, Frei C, Olma MH, Sumara G, Ricci R, et al. A Cul3-based E3 ligase removes Aurora B from mitotic chromosomes, regulating mitotic progression and completion of cytokinesis in human cells. *Dev Cell.* 2007;12:887–900. <https://doi.org/10.1016/j.devcel.2007.03.019>.
46. Ruzzenenti P, Asperti M, Mitola S, Crescini E, Maccarinelli F, Gryzik M, et al. The Ferritin-Heavy-Polypeptide-Like-17 (FTHL17) gene encodes a ferritin with low stability and no ferroxidase activity and with a partial nuclear localization. *Biochim Biophys Acta.* 1850;2015:1267–73. <https://doi.org/10.1016/j.bbagen.2015.02.016>.
47. Lim LS, Loh Y-H, Zhang W, Li Y, Chen X, Wang Y, et al. Zic3 is required for maintenance of pluripotency in embryonic stem cells. *Mol Biol Cell.* 2007;18:1348–58. <https://doi.org/10.1091/mbc.e06-07-0624>.
48. Lim LS, Hong FH, Kurnarso G, Stanton LW. The pluripotency regulator Zic3 is a direct activator of the Nanog promoter in ESCs. *Stem Cells.* 2010;28:1961–9. <https://doi.org/10.1002/stem.527>.
49. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature.* 2010;467:430–5. <https://doi.org/10.1038/nature09380>.
50. Yang S-H, Andrabi M, Biss R, Murtuza Baker S, Iqbal M, Sharrocks AD. ZIC3 controls the transition from naive to primed pluripotency. *Cell Rep.* 2019;27:3215–3227.e6. <https://doi.org/10.1016/j.celrep.2019.05.026>.
51. Kobayashi S, Fujihara Y, Mise N, Kaseda K, Abe K, Ishino F, et al. The X-linked imprinted gene family Fthl17 shows predominantly female expression following the two-cell stage in mouse embryos. *Nucleic Acids Res.* 2010;38:3672–81. <https://doi.org/10.1093/nar/gkq113>.
52. Argelaguet R, Clark SJ, Mohammed H, Stapel LC, Krueger C, Kapourani C-A, et al. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature.* 2019;576:487–91. <https://doi.org/10.1038/s41586-019-1825-8>.
53. Mohammed H, Hernando-Herraez I, Savino A, Scialdone A, Macaulay I, Mulas C, et al. Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. *Cell Rep.* 2017;20:1215–28. <https://doi.org/10.1016/j.celrep.2017.07.009>.
54. Cheng S, Pei Y, He L, Peng G, Reinius B, Tam PPL, et al. Single-cell RNA-Seq reveals cellular heterogeneity of pluripotency transition and X chromosome dynamics during early mouse development. *Cell Rep.* 2019;26:2593–2607.e3. <https://doi.org/10.1016/j.celrep.2019.02.031>.
55. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell.* 2014;159:635–46. <https://doi.org/10.1016/j.cell.2014.09.039>.
56. Heurtier V, Owens N, Gonzalez I, Mueller F, Proux C, Mornico D, et al. The molecular logic of Nanog-induced self-renewal in mouse embryonic stem cells. *Nat Commun.* 2019;10:1109. <https://doi.org/10.1038/s41467-019-09041-z>.
57. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One.* 2011;6:e19379. <https://doi.org/10.1371/journal.pone.0019379>.
58. Watanabe S, Umehara H, Murayama K, Okabe M, Kimura T, Nakano T. Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. *Oncogene.* 2006;25:2697–707. <https://doi.org/10.1038/sj.onc.1209307>.
59. Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, et al. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol.* 2011;13:838–45. <https://doi.org/10.1038/ncb2267>.
60. Dhanoa BS, Cogliati T, Satish AG, Bruford EA, Friedman JS. Update on the Kelch-like (KLHL) gene family. *Hum Genomics.* 2013;7:13. <https://doi.org/10.1186/1479-7364-7-13>.
61. Pintard L, Willems A, Peter M. Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J.* 2004;23:1681–7. <https://doi.org/10.1038/sj.emboj.7600186>.
62. Sowa ME, Bennett EJ, Gygi SP, Harper JW. Defining the human deubiquitinating enzyme interaction landscape. *Cell.* 2009;138:389–403. <https://doi.org/10.1016/j.cell.2009.04.042>.
63. Li Z, Fei T, Zhang J, Zhu G, Wang L, Lu D, et al. BMP4 Signaling Acts via dual-specificity phosphatase 9 to control ERK activity in mouse embryonic stem cells. *Cell Stem Cell.* 2012;10:171–82. <https://doi.org/10.1016/j.stem.2011.12.016>.

64. Oliviero G, Munawar N, Watson A, Streubel G, Manning G, Bardwell V, et al. The variant Polycomb Repressor Complex 1 component PCGF1 interacts with a pluripotency sub-network that includes DPPA4, a regulator of embryogenesis. *Sci Rep*. 2015;5:18388. <https://doi.org/10.1038/srep18388>.
65. Xie T, Pan S, Zheng H, Luo Z, Tembo KM, Jamal M, et al. PEG10 as an oncogene: expression regulatory mechanisms and role in tumor progression. *Cancer Cell Int*. 2018;18:112. <https://doi.org/10.1186/s12935-018-0610-3>.
66. Fonseca BD, Zakaria C, Jia J-J, Graber TE, Svitkin Y, Tahmasebi S, et al. La-related protein 1 (LARP1) represses terminal oligopyrimidine (TOP) mRNA translation downstream of mTOR complex 1 (mTORC1). *J Biol Chem*. 2015;290:15996–6020. <https://doi.org/10.1074/jbc.M114.621730>.
67. Ionomou M, Saunders DN. Systematic approaches to identify E3 ligase substrates. *Biochem J*. 2016;473:4083–101. <https://doi.org/10.1042/BCJ20160719>.
68. Haglund K, Dikic I. Ubiquitylation and cell signaling. *EMBO J*. 2005;24:3353–9. <https://doi.org/10.1038/sj.emboj.7600808>.
69. Parvatiyar K, Harhaj EW. Cell signaling. Anchors away for ubiquitin chains. *Science*. 2010;328:1244–5. <https://doi.org/10.1126/science.1192296>.
70. Declercq J, Sheshadri P, Verfaillie CM, Kumar A. Zic3 enhances the generation of mouse induced pluripotent stem cells. *Stem Cells Dev*. 2013;22:2017–25. <https://doi.org/10.1089/scd.2012.0651>.
71. Mutzel V, Okamoto I, Dunkel I, Saitou M, Giorgetti L, Heard E, et al. A symmetric toggle switch explains the onset of random X inactivation in different mammals. *Nat Struct Mol Biol*. 2019;26:350–60. <https://doi.org/10.1038/s41594-019-0214-1>.
72. Dorak MT, Karpuzoglu E. Gender differences in cancer susceptibility: an inadequately addressed issue. *Front Genet*. 2012; 3:268. <https://doi.org/10.3389/fgene.2012.00268>.
73. Chaligné R, Popova T, Mendoza-Parra M-A, Saleem M-AM, Gentien D, Ban K, et al. The inactive X chromosome is epigenetically unstable and transcriptionally labile in breast cancer. *Genome Res*. 2015;25:488–503. <https://doi.org/10.1101/gr.185926.114>.
74. Richardson AL, Wang ZC, De Nicolo A, Lu X, Brown M, Miron A, et al. X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell*. 2006;9:121–32. <https://doi.org/10.1016/j.ccr.2006.01.013>.
75. Vallot C, Ouimette J-F, Makhlouf M, Féraud O, Pontis J, Côme J, et al. Erosion of X chromosome inactivation in human pluripotent cells initiates with XACT coating and depends on a specific heterochromatin landscape. *Cell Stem Cell*. 2015;16:533–46. <https://doi.org/10.1016/j.stem.2015.03.016>.
76. Patel S, Bonora G, Sahakyan A, Kim R, Chronis C, Langerman J, et al. Human embryonic stem cells do not change their X inactivation status during differentiation. *Cell Rep*. 2017;18:54–67. <https://doi.org/10.1016/j.celrep.2016.11.054>.
77. Sahakyan A, Kim R, Chronis C, Sabri S, Bonora G, Theunissen TW, et al. Human naive pluripotent stem cells model X chromosome dampening and X inactivation. *Cell Stem Cell*. 2017;20:87–101. <https://doi.org/10.1016/j.stem.2016.10.006>.
78. Vallot C, Patrat C, Collier AJ, Huret C, Casanova M, Liyakat Ali TM, et al. XACT noncoding RNA competes with XIST in the control of X chromosome activity during human early development. *Cell Stem Cell*. 2017;20:102–11. <https://doi.org/10.1016/j.stem.2016.10.014>.
79. D'Antonio-Chronowska A, Donovan MKR, Young Greenwald WW, Nguyen JP, Fujita K, Hashem S, et al. Association of human iPSC gene signatures and X chromosome dosage with two distinct cardiac differentiation trajectories. *Stem Cell Reports*. 2019;13:924–38. <https://doi.org/10.1016/j.stemcr.2019.09.011>.
80. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015;517:583–8. <https://doi.org/10.1038/nature14136>.
81. Campeau E, Ruhl VE, Rodier F, Smith CL, Rahmsberg BL, Fuss JO, et al. A versatile viral system for expression and depletion of proteins in mammalian cells. *PLoS ONE*. 2009;4:e6529. <https://doi.org/10.1371/journal.pone.0006529>.
82. Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, et al. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol*. 2018;14:431–41. <https://doi.org/10.1038/s41589-018-0021-8>.
83. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31:827–32. <https://doi.org/10.1038/nbt.2647>.
84. Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *elife*. 2016;5 <https://doi.org/10.7554/eLife.19760>.
85. Concordet J-P, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*. 2018;46:W242–5. <https://doi.org/10.1093/nar/gky354>.
86. Sanson KR, Hanna RE, Hegde M, Donovan KF, Strand C, Sullender ME, et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat Commun*. 2018;9:5416. <https://doi.org/10.1038/s41467-018-07901-8>.
87. Morgens DW, Wainberg M, Boyle EA, Ursu O, Araya CL, Tsui CK, et al. Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nat Commun*. 2017;8:15178. <https://doi.org/10.1038/ncomms15178>.
88. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819–23. <https://doi.org/10.1126/science.1231143>.
89. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8:2281–308. <https://doi.org/10.1038/nprot.2013.143>.
90. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*. 2014;159:647–61. <https://doi.org/10.1016/j.cell.2014.09.029>.
91. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*. 2013;155:1479–91. <https://doi.org/10.1016/j.cell.2013.12.001>.
92. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014; 11:783–4. <https://doi.org/10.1038/nmeth.3047>.
93. Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc*. 2012;7:171–92. <https://doi.org/10.1038/nprot.2011.431>.
94. Matsuda T, Cepko CL. Controlled expression of transgenes introduced by in vivo electroporation. *Proc Natl Acad Sci USA*. 2007;104:1027–32. <https://doi.org/10.1073/pnas.0610155104>.
95. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9. <https://doi.org/10.1038/nmeth.1923>.
96. Gao Y, Xiong X, Wong S, Charles EJ, Lim WA, Qi LS. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat Methods*. 2016;13:1043–9. <https://doi.org/10.1038/nmeth.4042>.

97. Redolfi J, Zhan Y, Valdes-Quezada C, Kryzhanovska M, Guerreiro I, Iesmantavicius V, et al. DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat Struct Mol Biol.* 2019;26:471–80. <https://doi.org/10.1038/s41594-019-0231-0>.
98. Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 2005;33(Database issue):D501–4. <https://doi.org/10.1093/nar/gki025>.
99. Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, et al. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* 2014;42(Database issue):D756–63. <https://doi.org/10.1093/nar/gkt1114>.
100. Diaz AA, Qin H, Ramalho-Santos M, Song JS. HITSelect: a comprehensive tool for high-complexity-pooled screen analysis. *Nucleic Acids Res.* 2015;43:e16. <https://doi.org/10.1093/nar/gku1197>.
101. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* 2016;44:W3–10. <https://doi.org/10.1093/nar/gkw343>.
102. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* 2014;15:554. <https://doi.org/10.1186/s13059-014-0554-4>.
103. Li W, Köster J, Xu H, Chen C-H, Xiao T, Liu JS, et al. Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* 2015;16:281. <https://doi.org/10.1186/s13059-015-0843-6>.
104. Pilsner JR, Lazarus AL, Nam D-H, Letcher RJ, Sonne C, Dietz R, et al. Mercury-associated DNA hypomethylation in polar bear brains via the LUMInometric Methylation Assay: a sensitive method to study epigenetics in wildlife. *Mol Ecol.* 2010;19:307–14. <https://doi.org/10.1111/j.1365-294X.2009.04452.x>.
105. Chaumeil J, Augui S, Chow JC, Heard E. Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. *Methods Mol Biol.* 2008;463:297–308. https://doi.org/10.1007/978-1-59745-406-3_18.
106. Girardot C, Scholtalbers J, Sauer S, Su S-Y, Furlong EEM, Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics.* 2016;17:419. <https://doi.org/10.1186/s12859-016-1284-2>.
107. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
108. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010;26:589–95. <https://doi.org/10.1093/bioinformatics/btp698>.
109. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 2016;44:W160–5. <https://doi.org/10.1093/nar/gkw257>.
110. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
111. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26:139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
112. Hubner NC, Bird AW, Cox J, Splettstoesser B, Bandilla P, Poser I, et al. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J Cell Biol.* 2010;189:739–54. <https://doi.org/10.1083/jcb.200911091>.
113. Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods.* 2014;11:319–24. <https://doi.org/10.1038/nmeth.2834>.
114. Gielsch I, Meierhofer D. Metabolome and proteome profiling of complex I deficiency induced by rotenone. *J Proteome Res.* 2015;14:224–35. <https://doi.org/10.1021/pr500894v>.
115. Keilhauer EC, Hein MY, Mann M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol Cell Proteomics.* 2015;14:120–35. <https://doi.org/10.1074/mcp.M114.041012>.
116. Genolet O, Monaco AA, Dunkel I, Boettcher M, Schulz EG. Identification of X-chromosomal genes that drive sex differences in embryonic stem cells through a hierarchical CRISPR screening approach, Gene Expression Omnibus <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE142348>. (2021).
117. Genolet O, Monaco AA, Dunkel I, Boettcher M, Schulz EG. Identification of X-chromosomal genes that drive sex differences in embryonic stem cells through a hierarchical CRISPR screening approach, ProteomExchange <https://www.ebi.ac.uk/pride/archive/projects/PXD016729>. (2021).
118. Genolet O, Monaco AA, Dunkel I, Boettcher M, Schulz EG. Identification of X-chromosomal genes that drive sex differences in embryonic stem cells through a hierarchical CRISPR screening approach, ProteomExchange <https://www.ebi.ac.uk/pride/archive/projects/PXD017875>. (2021).
119. Argelaguet R, Clark SJ, Mohammed H, Stapel LC, Krueger C, Kapourani C-A, et al. Multi-omics profiling of mouse gastrulation at single-cell resolution. Github https://github.com/rargelaguet/scnmt_gastrulation. (2019).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.