EZH2 promotes neoplastic transformation through VAV interaction-dependent extranuclear mechanisms

Nandini Venkatesan[1,7], Jong Fu Wong[1,7], Kuan Pern Tan[2,6], Hwa Hwa Chung[1], Yin Hoe Yau[1], Engin Cukuroglu[5], Abdollah Allahverdi[1], Lars Nordenskiold[1], Jonathan Göke[5], Susana Geifman Shochat[1], Valerie Chun Ling Lin[1], M.S. Madhusudhan[2,3,4], I-hsin Su[1*]

[1]School of Biological Sciences, College of Science, Nanyang Technological University, 60 Nanyang Drive Singapore 637551, Republic of Singapore

[2]Bioinformatics Institute, 30 Biopolis Street, #07-01, Matrix, Singapore 138671, Republic of Singapore

[3]Indian Institute of Science Education and Research, Pune, India,

[4]Department of Biological Sciences, National University of Singapore, Singapore 117543

[5]Genome Institute of Singapore, 60 Biopolis Street, Genome, #02-01, Singapore 138672

[6]School of Computer Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

[7]These authors contributed equally to this work

*Correspondence should be addressed to: I-hsin Su (e-mail: ihsu@ntu.edu.sg)

Running title: **Extranuclear function of EZH2 in tumorigenesis**

KEY WORDS: EZH2, VAV, cytosolic function, adhesion dynamics, cellular transformation

ABSTRACT

Recently, we reported that the histone methyltransferase, EZH2, controls leukocyte migration through interaction with the cytoskeleton remodeling effector, VAV, and direct methylation of the cytoskeletal regulatory protein, Talin. However, it is unclear whether this extra-nuclear, epigenetic-independent function of EZH2 has a profound impact on the initiation of cellular transformation and metastasis. Here, we show that EZH2 increases Talin1 methylation and cleavage, thereby enhancing adhesion turnover and promoting accelerated tumorigenesis. This transforming capacity is abolished by targeted disruption of EZH2 interaction with VAV. Furthermore, our studies demonstrate that EZH2 in the cytoplasm is closely associated with cancer stem cell properties, and that overexpression of EZH2, a mutant EZH2 lacking its nuclear localization signal (EZH2ΔNLS), or a methyl-mimicking Talin1 mutant substantially promotes JAK2-dependent STAT3 activation and cellular transformation. Taken together, our results suggest a critical role for the VAV interaction-dependent, extra-nuclear action of EZH2 in neoplastic transformation.

INTRODUCTION

EZH2 overexpression or mutation is known to associate with several aggressive human solid tumor types, including prostate cancer, breast cancer, and different types of lymphomas, and is indicative of a poor prognosis in patients[37, 47, 66]. The oncogenic effect of EZH2 can be achieved through histone 3 lysine 27 trimethylation (H3K27Me3)-mediated transcriptional repression of tumor suppressor genes, such as CDH1 (E-cadherin), RKIP, RAD51, DAB2IP, RUNX3, and CDKN1c ($p57^{KIP2}$)[4, 6, 55, 74]. EZH2 can also regulate gene expression indirectly through the recruitment of DNA methyltransferases to target gene promoter regions[68]. PRC2-independent roles of EZH2 have also been identified. EZH2 interactions with NF-κB in breast cancer cells promote NF-κB target gene expression[40] and EZH2 recruitment to the *NOTCH1* promoter increases NOTCH1 expression and signaling[18]. While EZH2 acts as a coactivator for the androgen receptor (AR)[72] transcription factor, it can also enhance STAT3 activity or stabilize β-catenin through the methylation of these proteins[35, 80]. Furthermore, EZH2-mediated methylation of the transcription factors GATA4 and RORα results in the attenuation of their transcriptional activities[22, 39]. Interestingly, the oncogenic property of EZH2 has been challenged by several recent findings that suggest a tumor suppressor role for EZH2 in T-acute lymphoblastic leukemia[26, 52, 77]. Despite intensive efforts to determine the molecular mechanisms of dysregulated EZH2 expression in oncogenesis, we still do not fully understand the role of EZH2 in this context. Therefore, the underlying mechanisms linking dysregulated EZH2 expression with cancer initiation and progression may vary depending on the cellular context and oncogenic pathways that are activated in different cancers.

Functionally, EZH2 overexpression results in the down regulation of adhesion molecules, cell cycle regulators, transcription factors, and DNA repair machinery[4, 6, 10, 16, 55, 74], and these changes have been shown to contribute to cellular transformation and metastasis.

Furthermore, our previous study identified a cytosolic EZH2 complex associated with another group of proto-oncogenes, the VAV family of proteins[24, 60, 61]. VAV proteins are guanine nucleotide exchange factors (GEFs) for Rho family GTPases that are involved in cell adhesion, migration, and proliferation of normal cells as well as the pathogenesis of cancer cells[2, 3, 7, 33, 46]. More specifically, VAV2/VAV3 are shown to be essential for the development of breast cancer and skin tumors[12,46], while VAV2 is suggested to be recruited to epidermal growth factor receptor (EGFR) during the morphological transformation of mammary epithelial cells[64]. In addition, VAV1 gene mutations have been detected in various human tumors[2, 3, 33] and an oncogenic form of VAV1, which is N-terminally truncated and does not assume an auto-inhibitory conformation[1, 34], has been shown to promote the transformation of fibroblasts[19].

Even though there is a large amount of data showing the critical roles of EZH2 and VAV family proteins in tumorigenesis, it is not clear whether there is a functional link between EZH2 and VAVs in this context. Here, we identify the residues on EZH2 that are critical for its interaction with VAV and demonstrate that EZH2 interactions with VAV proteins are crucial for the regulation of adhesion dynamics and cellular transformation. Mechanistically, overexpression of wild-type EZH2 or EZH2 lacking its nuclear localization signal (EZH2□NLS) promotes Talin methylation and cleavage, thereby destabilizing adhesion structures in a VAV interaction-dependent manner. In addition, the expression of a methyl-mimicking Talin1 mutant or EZH2□NLS substantially enhances JAK2-dependent STAT3 activity and the expression of STAT3 target genes associated with tumorigenesis. Taken together, our data suggest that EZH2 interactions with VAV family proteins in the cytoplasm contribute to initial neoplastic transformation and possibly the maintenance of cancer stem cells through the regulation of adhesion dynamics and STAT3 signaling pathways, which precedes the well-established functions of EZH2 in epigenetic silencing of tumor suppressors.

RESULTS

Characterization and structural prediction of the EZH2-VAV1 interaction domains

The structural details of the EZH2-VAV interaction are required for determining whether it is critical for neoplastic transformation as well as for designing a treatment strategy to interrupt the interaction. With this in mind, we narrowed down the minimal VAV1 interaction domain to $EZH2_{201-252}$ and confirmed the direct interaction of EZH2-VAV in the absence of other complex proteins by reciprocal pull-down experiments using recombinant EZH2 ($EZH2_{201-252}$) and VAV1 ($VAV1_{1-172}$)[24, 25] fragments (**Fig. 1a**). The binding affinity of the EZH2 and VAV1 fragments (**Fig. 1b**) was further determined using the surface plasmon resonance (SPR)-based biosensor, BIACORE, to be between 0.6 to 1.5 μM with a stoichiometry of 1:1 (**Fig. 1c, d**). This interaction is characterized as a quick interaction with a binding affinity similar to other important biological interactions[57].

In order to better understand the EZH2-VAV interaction on a molecular level, we attempted to co-crystallize the interaction domains of these two proteins. Unfortunately, all of our crystallization attempts failed, so we decided to employ a bioinformatics approach as an alternative to determine the critical amino acid residues required for this interaction. Since the minimal EZH2-interacting domain of VAV1 (amino acids 66-115, reported in Ref[24, 25]) contains a calponin homology (CH) domain, we examined all 8 known protein-protein interactions involving CH domains and observed that they all interact with helices. This pattern of CH domain interaction with helices was consistent even when the sequences of interacting helices were only weakly similar to one another (for details, see Supplementary Methods). Furthermore, electrostatic profiling of the VAV1 CH domain shows a charged interaction surface (**Fig. 1e**) where one part of the surface is positively charged and the other end is negatively charged. Based on these analyses, we propose that the part of EZH2 interacting

with VAV1 must be helical and that the interaction is likely to be predominated by electrostatic interactions, which is a conserved feature of all observed interactions involving CH domains.

On EZH2, two helices are predicted between residues 221-230 and residues 237-250 (**Supplementary Fig. 1a**). Interestingly, a similar structure has been reported recently in EZH2's yeast counterpart, despite very low sequence homology between the yeast and mammalian EZH2 proteins[30]. Our analyses indicate that charge complementarity between VAV1 and EZH2 is best achieved with the second helix and thus the interaction models were fine tuned to maximize the charge complementarity. The detailed structural analysis is shown in the Supplementary Methods.

Identification of the critical residues required for EZH2-VAV1 interaction

Following our structural analyses, we generated a model of the VAV1-EZH2 complex (**Fig. 1e**), proposing that the positively charged amino acid residues of the EZH2 helix (K234, K241 and K245) complement the negatively charged amino acid residues of VAV1 (E92, E95, D98, D101 and D104) (**Fig. 1f** and **Supplementary Methods**). The proposed EZH2-interaction region is highly conserved between VAV1, VAV2, and VAV3 from various species (**Fig. 1g** and **Supplementary Fig. 1b**), suggesting that the EZH2-VAV interaction is likely to be conserved amongst different cell types and species. Consistent with this notion, EZH2 was shown to interact with VAV2 in non-hematopoietic lineage cells such as BOSC23, MCF10A cells, and a human breast cancer cell line (**Supplementary Fig. 2a**).

To validate our predictions, the proposed mutations were introduced into the recombinant $EZH2_{201-252}$ fragment and the resulting mutants were then tested for their abilities to interact with the wild-type VAV1 fragment using pull-down assays, as described in **Figure 1a**. Our data revealed that alanine substitution of lysine residues K234, K241 and K245 on

EZH2 resulted in a modest, but consistent, reduction of EZH2-VAV1 interaction (**Fig. 1h**). Subsequently, we generated the double lysine mutants EZH2-K241/K245A (reported in[20]) and EZH2-K234/K245A, which exhibited significant decreases in VAV1 interaction in our *in vitro* assays (**Fig. 1h**). Using SPR analysis, these EZH2 mutants were shown to bind to VAV1 fragments with 5 to 10-fold reduced affinities (**Fig. 1i**). Since EZH2-K234/K245A exhibited more reductions in VAV1 interaction in our assays, we utilized the EZH2-K234/K245A mutant in our subsequent studies (hereafter referred to as EZH2-VAV/MT). The reduced interaction of mutated EZH2 and VAV1 was further verified using full-length proteins (**Supplementary Fig. 2b**) and the additional K241A mutation did not further reduce the interaction (**Supplementary Fig. 2c**). Similarly, complementary mutations in full-length VAV1 (E92A, E98A, and D104A) also led to substantial reductions in their interactions with EZH2 (**Supplementary Fig. 2d**). Moreover, we confirmed that the interaction of EZH2-VAV/MT with VAV2, the major VAV family protein in non-hematopoietic lineage cells, was also significantly reduced both *in vitro* (**Fig. 1j)** and in cells (**Supplementary Fig. 2e)**. However, these lysine mutations on EZH2 did not disrupt interactions with components of the PRC2 core complex (SUZ12 and EED) (**Supplementary Fig. 2f)** or compromise EZH2's histone methyltransferase (HMTase) activity *in vitro* (**Fig. 1k, l**).

VAV interaction is required for EZH2-promoted tumorigenesis

We next sought to determine whether the interaction of EZH2 with VAV family proteins is essential for the transforming capacity of EZH2. For our initial *in vitro* assessment of transforming capacity, we overexpressed wild-type EZH2, an enzymatically inactive form of EZH2 (EZH2H689A), or the EZH2-VAV/MT mutant in the mammary epithelial cell line H16N2 and tested whether they could initiate malignant transformation using anchorage-independent cell growth assays. To enhance the EZH2 interaction with VAV in the cytosol, we

also expressed a mutant EZH2 lacking its nuclear localization signal (EZH2ΔNLS). To avoid interference, endogenous expression of EZH2 in H16N2 cells was knocked down by lentivirus-mediated shRNA and the target site of the shRNA was mutated in the exogenous EZH2 constructs (**Supplementary Fig. 3a**). The total amounts of exogenous EZH2 variants were about 4 to 5-fold higher than those observed in normal epithelial cells, but were physiological and comparable to the endogenous EZH2 expression levels in 4T1 and MDA-MB-231 breast cancer cells (**Supplementary Fig. 3a**). The results of our anchorage-independent cell growth assays showed that the expression of wild-type EZH2 promoted colony formation (**Fig. 2a**), which is consistent with previous studies[6, 37]. Furthermore, we found that expression of the EZH2ΔNLS mutant resulted in a 2-fold increase in cytosolic EZH2 expression levels (**Supplementary Fig. 3b**), drastically elevated the number of breast cancer stem-like cells (**Supplementary Fig. 3c**), and significantly promoted cellular transformation relative to results obtained for wild-type EZH2 (**Fig. 2a**, left and middle panels). In contrast, expression of the enzymatically inactive EZH2H689A mutant or control construct (MIG) did not promote anchorage-independent cell growth (**Fig. 2a**). Moreover, VAV2 knockdown drastically reduced the transforming capacity of EZH2ΔNLS (**Fig. 2a**, right panel, and **Supplementary Fig. 3d**). Similarly, expression of either the EZH2-VAV/MT mutant or an EZH2-VAV/MT mutant lacking its NLS resulted in the complete loss of EZH2-promoted cellular transformation (**Fig. 2a**).

To determine whether VAV interaction is crucial for EZH2-promoted primary tumor growth *in vivo*, we first utilized a xenograft model of SCID mice transplanted with human breast cancer cells. For our experiments, EZH2 variants were transduced into MDA-MB-231 cells and the resulting cells were inoculated into the mammary fat pads of female NOD-SCID mice (expression of EZH2 variants is shown in **Supplementary Fig. 3e**). Even though the

percentage of exogenous EZH2 in these cells only ranged from 30-45% (**Supplementary Fig. 3e**), we still observed distinct influences of EZH2 variants on tumor growth *in vivo*. The tumor growth in mice that received wild-type EZH2- or EZH2ΔNLS-expressing cells was significantly higher than the growth observed in mice that received cells expressing MIG, EZH2H689A, or EZH2-VAV/MT (**Fig. 2b, c**). Mice that received cells expressing wild-type EZH2 or EZH2ΔNLS also showed enlarged lymph nodes (LNs), which is indicative of LN metastasis (**Fig. 2d** and **Supplementary Fig. 3f** showing the presence of human origin cancer cells in LNs).

To further validate the oncogenic capacity of EZH2 variants *in vivo* in a different experimental setting, we utilized the 4T1 mouse breast cancer model. Endogenous EZH2 in 4T1 cells was first lowered by specific shRNA-mediated knockdown before the cells were transduced with EZH2 mutants and inoculated into the mammary fat pads of BALB/c mice (the expression of EZH2 variants is shown in **Supplementary Fig. 3g**). In these experiments, the exogenous EZH2 consisted of 40-65% of the total EZH2 (**Supplementary Fig. 3g**, right panel). It was not possible to drastically increase the expression of exogenous EZH2 in cancer cells, likely due to the high level of endogenous EZH2 or potentially oncogene toxicity[21]. Our results showed that even under immune-competent conditions, EZH2 and EZH2ΔNLS were still able to greatly enhance tumor growth (**Fig. 2e-g**). Interestingly, when we reduced the VAV-targeting capacity of the EZH2 lacking its NLS (EZH2-VAV/MTΔNLS), the elevated tumor growth and metastasis were reduced, demonstrating the critical role of the EZH2-VAV interaction in EZH2-promoted tumorigenesis (**Fig. 2e-g**).

EZH2 promotes VAV interaction-dependent adhesion turnover

Previous work from our lab has shown that extra-nuclear EZH2 has important roles in regulating receptor-induced actin polymerization and integrin-dependent transendothelial

migration in a number of different cell types[20, 60]. Since malignant transformation is frequently associated with changes in the adhesive and migratory behaviour of cells, we hypothesized that in addition to the epigenetic regulation of gene expression, EZH2 also controls adhesion dynamics and thereby contributes to tumorigenesis in a VAV-dependent manner. To test our hypothesis, we first investigated the formation of adhesion structures in H16N2 cells expressing EZH2 variants. While the adhesion assembly rates were comparable between cells expressing EZH2 variants, the expression of wild-type EZH2 or EZH2 lacking its NLS substantially reduced the size (**Fig. 3a** and **3b**, top panel) and number (**Fig. 3a** and **3b**, bottom panel) of adhesion structures by promoting adhesion disassembly (**Fig. 3c**, **Supplementary Video 1**). Correspondingly, the areas occupied by these cells were significantly reduced (**Fig. 3d**). In contrast, the expression of inactive EZH2 (EZH2H689A) or VAV interaction mutant EZH2 (EZH2-VAV/MT) in H16N2 cells significantly promoted the formation of large, stable, adhesion structures, such as focal and fibrillar adhesions, with very slow disassembly rates (**Fig. 3a-c** and **Supplementary Video 2**), resulting in cells that spread out extensively (**Fig. 3d**). Similarly, even in cells expressing hyperactive EZH2 (EZH2Y641F), when the interaction with VAV was disrupted, the reduction in cell spreading area was abolished, (**Fig. 3d**, EZH2-VAV/MT-Y641F). Overall, these results are well correlated with the oncogenic capacity of the EZH2 variants shown earlier and suggest that EZH2 promotes cellular transformation at least partly through the regulation of adhesion dynamics.

EZH2-mediated Talin1 methylation regulates EGF-independent adhesion dynamics

Since VAV2 is known to be recruited to epidermal growth factor receptor (EGFR) and >30% of all breast cancers are linked to overexpression of EGFR or abnormal EGFR signaling, we sought to determine whether EGF induced disassembly of adhesions is affected by the expression of EZH2 variants. For these experiments, the aforementioned EZH2

knockdown H16N2 cells expressing EZH2 variants were EGF starved for 3 days and then re-stimulated with EGF. This treatment significantly promoted the disassembly of adhesion structures in EZH2 knockdown cells and cells expressing wild-type or cytosolic EZH2 (**Supplementary Fig. 4**). However, the prominent adhesion structures in cells expressing inactive or VAV-targeting mutant EZH2 were less sensitive to EGF induced disassembly (**Supplementary Fig. 4)**. Moreover, EGFR signaling events, such as activation of EGFR, AKT, FAK, ERK, and the VAV downstream effector PAK1 (**Supplementary Fig. 5a, b**), as well as recruitment of VAV2 to EGFR, VAV2 phosphorylation, and EGFR degradation[64] (**Supplementary Fig. 5c-e**) were not altered in EZH2 variants expressing mammary epithelial cells. These data suggest that even though the expression of EZH2 reduced the formation of prominent adhesion structures, EGF induced disassembly of focal complexes/adhesions and receptor proximal signaling are likely to be independent of EZH2.

Interestingly, we repeatedly observed an increased abundance of cleaved Talin1 rod and head domains in cells expressing wild-type EZH2 or EZH2ΔNLS, but not in cells expressing inactive or VAV interaction mutant EZH2 (**Fig. 4a**). Similar results for EZH2-promoted Talin1 cleavage were observed in BOSC23 cells expressing EZH2 variants (**Fig. 4b**). Since our recent study in innate leukocytes revealed that EZH2-mediated methylation of Talin1 at lysine residue 2454 regulates cell migration[20] and similar EZH2-promoted Talin1 methylation (**Fig. 4c**) and EZH2-VAV-Talin1 complexes (**Fig. 4d**) were also observed in mammary epithelium and breast cancer cells, we hypothesized that EZH2-mediated Talin1 methylation first interferes with F-actin binding, then promotes Talin1 cleavage and subsequent disassembly of adhesion structures. Indeed, when we expressed a mutant Talin1 in which the EZH2-methylated Lys2454 was mutated to an un-methylatable alanine residue or an un-methylatable, polar glutamine residue, the binding to F-actin remained the same (K2454A, **Fig.**

4e, left) or was increased (K2454Q, **Fig. 4e**, right) and cleavage of these Talin1 mutants in cells was reduced (**Fig. 4f**). Conversely, the replacement of Talin1-Lys2454 with a tri-methyl-mimicking phenylalanine residue[28] (K2454F) resulted in decreased F-actin binding (**Fig. 4e**) and enhanced cleavage (**Fig. 4f**). Moreover, EZH2-promted Talin cleavage was dependent on the presence of VAV2 (**Fig. 4g**). The effects of EZH2-mediated Talin1-K2454 methylation on adhesion structures were further confirmed in H16N2 cells expressing the Talin1-K2454 mutants described above. The expression of Talin1-K2454A or -K2454Q in H16N2 cells promoted the formation of stable focal and fibrillar adhesions (**Fig. 4h** and **Supplementary Fig. 5f**) with extremely slow disassembly rates (**Fig. 4i** and **Supplementary Video 3**), whereas expression of Talin1-K2454F resulted in rapid disassembly of adhesions (**Fig. 4i** and **Supplementary Video 4**) and increased numbers of H16N2 cells without visible adhesion structures (**Supplementary Fig. 5g**). These results suggest that EZH2-mediated Talin1 methylation promotes Talin1 cleavage and subsequently leads to disassembly of adhesion structures.

EZH2 lacking its NLS modulates gene expression profiles favoring tumorigenesis.

In order to evaluate the effects of EZH2 mutation on nuclear histone methylation, we examined the overall H3K27 trimethylation (H3K27me3) levels in H16N2 cells expressing EZH2 variants. As expected, the H3K27me3 levels in cells expressing wild-type-EZH2 or VAV interaction mutant EZH2 were comparable (**Fig. 5a**), confirming the unaltered histone methylation capacity of the VAV targeting mutant EZH2 compared to that of wild-type EZH2, as shown earlier (**Fig. 1k, l**). In contrast, the H3K27me3 level was reduced in cells expressing EZH2H689A reflecting the reduced enzymatic activity (**Fig. 5a**). Even though we showed that deletion of the nuclear localization signal in EZH2 does not affect the interaction of EZH2 with PRC2 core components, EZH2's *in vitro* HMTase activity (**Fig. 5b**, **Supplementary Fig. 6**,

Fig. 1k, and Ref[60]), or the capacity of EZH2 to methylate cytosolic Talin1 in cells (**Fig. 4c**), we found that overexpressed EZH2ΔNLS in the nucleus (**Supplementary Fig. 3b**) contributes little to H3K27 methylation (**Fig. 5a**). This is due to the disrupted interaction of EZH2ΔNLS and CDYL (**Fig. 5c**, top), a co-factor that is required for chromatin targeting and efficient H3K27 methylation in cells[78]. Consequently, the recruitment of EZH2ΔNLS to EZH2 target genes, such as *ATF3* or *KDR*, was reduced (**Fig. 5c**, middle and bottom panels). The increased EZH2 levels in the cytosol may have also sequestered other components of the PRC2 complex in the cytoplasm and reduced the histone methyltransferase capacity in the nucleus. However, the distributions of other members of the PRC2 complex, including SUZ12, EED, and HDAC1, in the nucleus were not affected by the expression of any exogenous EZH2 variants (**Fig. 5d**), but reduced SUZ12 and EED protein levels in EZH2 knockdown cells was observed, as reported previously[5, 53].

Interestingly, the expressions of several well-established transcriptional target genes of EZH2 in cancer cells (such as E-cadherin, *RKIP*, and *RAD51*) were comparable in EZH2-knockdown cells or cells expressing EZH2 variants (**Fig. 6a**), suggesting that their down-regulation in cancer cells is a consequence rather than a cause for cellular transformation and that their expressions are not regulated by EZH2 in non-cancerous H16N2 cells when cultured on 2-D surfaces. Similarly, when we further determined the molecular changes that occur in H16N2 cells expressing EZH2 variants by performing whole-transcriptome analyses with RNA sequencing (RNA-seq), the overall gene expression patterns were not much influenced by EZH2 knockdown, the expression of EZH2 variants, or changes in H3K27me3 levels. There were also only a small number of differentially expressed genes (DEGs) in EZH2 knockdown cells (MIG) or cells expressing EZH2H689A or EZH2-VAV/MT, compared to those observed in wild-type EZH2-overexpressing cells (**Supplementary Fig. 7a, b** and source data for Fig.

6). In contrast, EZH2ΔNLS-expressing cells displayed a very different gene expression profile (**Supplementary Fig. 7a, b** and source data for Fig. 6). Using the ENCODE-motifs online tool to analyze the promoters of DEGs in EZH2ΔNLS-expressing cells, we found that EZH2 target genes were highly enriched among up-regulated, differentially expressed genes (23%, **Fig. 6b**, $P = 2.3 \times 10^{-11}$, Fisher's exact test, **Fig. 6c**), which correlates with the diminished H3K27me3 level in these cells. However, this high enrichment score was only achieved when we took all EZH2 target genes from available ChIP data in various cell types into account. If we only considered EZH2 targets in epithelial cells, the enrichment was substantially reduced (8%, $P = 0.0006$, Fisher's exact test, **Fig. 6c**, red bar). Taken together, these results suggest that EZH2ΔNLS-expressing cells may be losing the molecular signature of epithelial cells. Moreover, most of these up-regulated genes were not de-repressed in EZH2H689A-expressing cells or EZH2-knockdown cells (source data for Fig. 6), suggesting that EZH2ΔNLS may promote their expression through additional mechanisms other than simply alleviating H3K27-mediated repression.

Additionally, we found that a remarkably large fraction of DEGs in EZH2ΔNLS-expressing cells were highly enriched for STAT3 binding sites in their promoters (26% of up-regulated genes, $P = 1.4 \times 10^{-6}$; 31% of down-regulated genes, $P = 3.8 \times 10^{-14}$, Fisher's exact test, **Fig. 6b, c**). Since we did not observe an increase in STAT3 mRNA or protein levels (source data for Fig. 6 and **Fig. 7a-c**), these results indicate elevated STAT3 activities in EZH2ΔNLS-expressing cells. Furthermore, several STAT3 target genes, including the transcription factors *STAT1*, *STAT2*, *IRF1*, and *c-Fos*, were up-regulated in EZH2ΔNLS-expressing H16N2 cells (source data for Fig. 6). Therefore, the expressions of target genes of these transcription factors were altered and highly enriched among the DEGs as well (31% of up-regulated genes, Fisher's exact test, $P = 2.9 \times 10^{-9}$; 35% of down-regulated genes, Fisher's

exact test, P = 5.5 x 10^{-17}, **Fig. 6b, c**). However, the prominent oncogene, *MYC*, was down-regulated, while several MYC inhibitors, such as *Id2*, *Mxi1*, and *Mxd4*, were upregulated in cells expressing EZH2ΔNLS (source data for Fig. 6)[41, 56, 71]. Consequently, most of the down-regulated genes are known MYC target genes (73%, P = 4.1 x 10^{-50}, Fisher's exact test, **Fig. 6b, c**). Even though the downregulation of *MYC* may appear to be counterintuitive given the elevated oncogenic capacity of EZH2ΔNLS, the expression of a constitutively active STAT3 mutant in mesenchymal stem cells has also been shown to down-regulate MYC[50]. Reduced *MYC* expression may have also directly resulted in the observed down-regulation of several tumor suppressors, such as *BRCA1*, *XRCC2*, *NF2*, *FH*, and *RASSF1*, in EZH2ΔNLS-expressing cells (source data for Fig. 6).

EZH2 promotes STAT3 activity and cellular transformation through methylation of Talin1

The increased activities of STAT family proteins and the reduced expressions of tumor suppressors are associated with cancer cell survival and growth[75] and are therefore likely to contribute, in large part, to the elevated oncogenic activity of EZH2ΔNLS. Since overexpression of EZH2 is reported to activate STAT3 through methylation[13, 35], we wondered whether EZH2-mediated STAT3 methylation is dependent on its interaction with VAV. Surprisingly, we found that EZH2-mediated STAT3 methylation did not show obvious changes in cells expressing EZH2 variants (**Fig. 7a**, with quantification shown in **Supplementary Fig. 7c**), suggesting that STAT3 methylation is unlikely to play a role in this context. The residual STAT3 methylation observed in EZH2H689A-expressing cells is expected (**Fig. 7a**), since STAT3 is also methylatable by SET9[73]. However, phosphorylation of STAT3 on Tyr 705 (pY705) by JAK2, which is required for STAT3 dimerization and activation, was increased in cells expressing wild-type EZH2 and substantially elevated in cells expressing EZH2ΔNLS,

while phosphorylation of this site was diminished by the disruption of EZH2 interaction with VAV (**Fig. 7b** and **Supplementary Fig. 7d**). In contrast, phosphorylation of Ser 727 (pS727) on STAT3 by ERK MAP kinase[11] was not affected by the expression of EZH2 variants (**Fig. 7b**). Furthermore, VAV interaction was specifically required for EZH2 or cytosolic EZH2-promoted phosphorylation of Y705, but was dispensable for IL-6-stimulated STAT3 activation (**Fig. 7c**), suggesting a cytokine-independent, novel STAT3 activation signaling pathway mediated by cytosolic EZH2.

Since interaction with VAV is also required for EZH2-mediated Talin methylation, we sought to determine whether the methylation state of Talin1 could also have an effect on STAT3 activation. Indeed, expression of the tri-methyl-mimicking Talin1 mutant (K2454F) was sufficient to stimulate substantial STAT3 phosphorylation (pY705) and activation (**Fig. 7d** and **Supplementary Fig. 7e,f**). Moreover, EZH2-, EZH2ΔNLS-, and methyl-Talin mutant-promoted STAT3 phosphorylation was dependent on JAK2 activity (**Fig. 7b** and **Supplementary Fig. 7f**). Consequently, similar to EZH2 and EZH2ΔNLS, the methyl-mimicking Talin1 mutant also significantly promoted cellular transformation (**Fig. 7e**) that was dependent on JAK/STAT activity (**Fig. 7f)**. Furthermore, the expressions of several key factors that were up-regulated in EZH2ΔNLS-expressing cells were also up-regulated by methyl-mimicking Talin1 (**Fig. 7g**). Overall, our data suggest that cytosolic EZH2 contributes to the initiation of cellular transformation or the formation of cancer stem cells through VAV-dependent Talin methylation and STAT3 activation.

EZH2 is enriched in the cytoplasm of human breast cancer stem cells

To validate the physiological relevance of cytosolic EZH2 in human breast cancer cells, we first sorted MDA-MB-231 human triple negative breast cancer cells based on their surface expression levels of CD44 and their capacity to efflux Rhodamine 123 dye, and then analyzed

the sub-cellular distribution of EZH2 in the sorted cells. We found that EZH2 was significantly enriched in the cytoplasmic compartment of $CD44^{High}$ Rhodamine 123 negative cancer stem cells with the side population phenotype ($CD44^{Hi}SP$) (**Fig. 8a-c** and **Supplementary Video 5**), whereas cytosolic EZH2 was largely depleted in $CD44^{Low}$ Rhodamine 123 positive cells ($CD44^{Lo}Rd^{+}$) (**Fig. 8a-c** and **Supplementary Video 6**). Correspondingly, $CD44^{Hi}SP$ cells exhibited higher STAT3 activity than $CD44^{Lo}Rd^{+}$ cells (**Fig. 8c**). Taken together, our results suggest a cytosolic role for EZH2 in cellular transformation, likely through methylation of Talin and the activation of STAT3 in human breast cancer stem cells. It is also possible that cytosolic EZH2 contributes to the generation of tumor-initiating cells partly by amending their adhesion properties

DISCUSSION

Here, we report that the polycomb group protein, EZH2, promotes cellular transformation and metastasis through VAV interaction-dependent mechanisms. While EZH2 lacking its NLS exhibited much greater transforming capacity than wild-type EZH2, targeted disruption of EZH2 interaction with VAV abolished both wild-type EZH2- and EZH2ΔNLS-promoted cellular transformation. Because the well-established VAV effector functions are relatively unaffected by the absence of EZH2 or the expression of EZH2 variants, we believe that VAV is likely acting as an adaptor in this context. We have shown that EZH2-VAV interaction is critical for EZH2's recruitment to Talin and subsequent Talin methylation-promoted adhesion turnover. Furthermore, EZH2-VAV interaction may also bring JAK2 into close proximity with EZH2-interacting STAT3, thereby promoting phosphorylation of STAT3. We hypothesize that both of these VAV interaction-dependent events (adhesion turnover and STAT3 activation) are critical for EZH2-mediated tumorigenic signaling (see model in Supplementary Fig. 8).

Mechanistically, EZH2 is not involved in the EGF signaling pathway, a key factor in most epithelial malignancies[51]. EGF-stimulated VAV2 activation and down-stream signaling events, including activation of VAV effectors, were not affected by EZH2 knockdown or overexpression of EZH2 variants. Moreover, even though cytosolic EZH2-expressing cells possess high oncogenic capacity, the expression of several well-characterized tumor suppressors that are repressed by EZH2 in cancer cells were not regulated by EZH2 in mammary epithelial cells. Instead, our data suggest that EZH2-VAV2 interaction controls adhesion turnover and cellular transformation through Talin1 methylation and cleavage. In fact, elevated focal adhesion remodeling of cells *in vitro* is frequently correlated with elevated migratory rate and metastatic capacity *in vivo*. Furthermore, over 90% of all cancers originate

from mutations in epithelial cells[63], therefore, changes in adhesion properties of these cells are particularly critical for tumor initiation, progression, and metastasis[43, 48]. In this context, EZH2 and methylated Talin specifically regulate adhesion turnover, but do not affect adhesion formation. These changes in adhesion properties may also directly contribute to the oncogenic capacity of epithelial cells by amending their gene expression patterns[29, 67].

Interestingly, aside from altered adhesion properties, STAT3 activity in cytosolic EZH2-expressing cells was also significantly elevated. Members of the STAT family of proteins are known to transiently regulate gene expression in response to cytokines and growth factors. Constitutively activated STAT3 is also found in more than 50% of primary breast tumors and tumor-derived cell lines[27]. In addition to the direct contribution of activated STAT3 to the elevated oncogenic capacity of EZH2ΔNLS-expressing cells, subsequent upregulation of STAT1/2 by STAT3 is likely to be responsible for the increased expressions of various claudins and MMPs that are associated with heightened metastatic capabilities[17, 54].

Our data also revealed that JAK2 activity was critical for EZH2- and methyl-mimicking Talin mutant-promoted STAT3 phosphorylation (pY705) and activation. JAK2 has been reported to interact with VAV1 in myeloid cells[44] and may also associate with VAV2 in epithelium or breast cancer cells to implement EZH2-promoted STAT3 activation. Indeed, EZH2's capacity to interact with VAV was found to be essential for EZH2-, but not IL-6-promoted STAT3 phosphorylation, suggesting a specific, novel role for VAV in EZH2-promoted STAT3 activation during cellular transformation. Interestingly, EZH2-mediated STAT3 methylation is not dependent on EZH2's interaction with VAV2. However, the two reported EZH2-mediated STAT3 methylation sites[13, 35] are likely to fine tune the conformation of STAT3, making it more accessible for phosphorylation and final activation, as demonstrated by reduced STAT3 (pY705) in EZH2H689A-expressing cells. Overall, elevated STAT3

signaling is likely to be critical for the enhanced transforming capacity of EZH2 and cytosolic EZH2 in our experimental setting.

In addition, genes implicated in promoting epithelial-mesenchymal transition (EMT) or those expressed in cancer stem cells, such as *CD44*, various *ALDHs*, *KLF8*, *S1PR1*, *Bmi1*, *CCL5*, *ZEB2*, and *SNAI1* were up-regulated in cells expressing EZH2ΔNLS. We also found that cytoplasmic EZH2 was enriched in human breast cancer stem cells. Consistent with our results, abundant EZH2 in the cytosol has been reported to associate with primary human glioma stem-like cells[49]. Collectively, cytosolic EZH2-promoted adhesion turnover through Talin1-methylation and STAT3 activation may promote the initiation of EMT and cancer stem cell formation before the establishment of oncogenic transcriptional programs by nuclear EZH2.

The enormous effort that has been put into developing EZH2-specific methyltransferase inhibitors as anticancer drugs, including DZNep[62], GSK126 (Ref[45]), and the specific compound, EPZ005687 (Ref[38]), have yielded promising results in laboratory or pre-clinical settings[9, 36] and several specific drugs targeting EZH2 and other epigenetic regulators are currently in clinical trials[65, 76]. However, the clinical application of these drugs should be tested with caution. Our previous studies have shown that EZH2 plays indisputable roles in lymphocyte development and function[59, 60] and several studies also reveal that EZH2 is crucial for the maintenance of various somatic stem cells, including skin, muscle, hematopoietic and neural stem cells[8, 16, 31, 32, 42]. Additionally, tumor suppressor and methyltransferase-independent functions of EZH2 in tumorigenesis have been suggested recently[26, 40, 52, 77]. Prolonged EZH2 depletion has also been recently reported to lead to an epigenetic switch in cancer cells that results in enhanced tumor progression[15]. Similarly, deletion of *Ezh2* gene in mice could result in genome instability making Ezh2-deficient cells more susceptible for other oncogene induced tumorigenesis[69]. Moreover, treatment with an EZH2 inhibitor, such as

DZNep or GSK126, has been proven ineffective on glioma stem-like cells[49] and has been shown to restrict anti-tumor immunity of cytotoxic T cells[79]. For these reasons, EZH2-specific methyltransferase inhibitor-based cancer treatments may not be the best strategy to combat EZH2-associated cancers.

In summary, the data reported here demonstrate that the polycomb group protein, EZH2, in addition to its well-established roles in epigenetic silencing of tumor suppressors[14, 23, 58], also controls cellular transformation and tumor growth through VAV interaction-dependent Talin1 methylation. Nuclear functions of EZH2 are frequently involved in regulating basic cell survival and proliferation, whereas cytosolic pathways are more specific for fine tuning protein functions in certain physiological or pathological contexts. By targeting the extra-nuclear functions of EZH2 through the disruption of the EZH2-VAV interaction, we may be able to effectively control cancer progression without eliciting serious resistance. In the clinical setting, disruption of EZH2 interaction with VAV in circulating and disseminated tumor cells may result in migratory arrest of the tumor cells, similar to results observed in Ezh2-deficient leukocytes, by forming abnormally stable adhesions with the microvascular wall[20] and rendering them unable to transmigrate into interstitial tissues and establish metastases. Thus, our findings not only identify a unique regulatory role of EZH2 in tumorigenesis, but also provide a mechanistic basis for developing novel therapeutic intervention strategies in the treatment of human cancers associated with EZH2 overexpression or drug resistant stem-like tumour cells.

EXPERIMENTAL PROCEDURES

Recombinant protein production

VAV_{1-172} was cloned into PGEX-4T and $EZH2_{201-252}$ was cloned into pET-21b. The VAV1 fragment was derived from human cDNA. The EZH2 fragment was derived from mouse cDNA, however, the amino acid sequence of the relevant domain is highly conserved between human and mouse. They were grown and expressed in BL21 (DE3) cells to an OD_{600} of 0.6 at 37℃ and induced with 1mM IPTG for 3 hrs. GST-VAV was purified using Glutathione Sepharose beads and EZH2-His was purified using His-Pur Cobalt resin. Individual mutations in $EZH2_{201-252}$ were created using the Stratagene Quik change site-directed mutagenesis using primers as described and were then purified.

Cell culture and plasmids

HEK-293 cells used for expression of full-length EZH2 were cultured in DMEM supplemented with 10% FCS. Jurkat T-cells were grown in RPMI medium containing 10% FCS and were used for immunoprecipitation of endogenous VAV1. H16N2 and MCF10A mammary epithelial cells lines were grown in supplemented F-12 or DMEM media. Cells were transfected with MSCV-IRES-GFP (MIG) vector based retroviruses encoding EZH2, EZH2 lacking its nuclear localization signal-RKKKRK (EZH2ΔNLS), enzymatically inactive EZH2 (EZH2H689A), or the EZH2-VAV interaction mutant (EZH2-VAV/MT or EZH2-VAV/MTΔNLS). Cells were sorted based on their expression of GFP by FACS-ARIA Endogenous EZH2 was knocked down using shRNA against *EZH2* (V2LHS_17509, Open Biosystems) and selected with puromycin (1 μg/ml). The shRNA target nucleotides on *EZH2* cDNA were mutated (1910-5’ tctcagaatactgtgggga**c**at**g**atttctcaggatgaagc**a**g 3’1951, bold underlined letters are mutated.). MDA-MB-231 cells for mammary tumor xenografts were

grown in DMEM supplemented with 7.5% FCS. Cells expressing EZH2 or variants (as described) were obtained using retrovirus-mediated expression (as described) followed by cell sorting based on their GFP expression using FACS-ARIA.

Immunoprecipitation

EZH2-HIS $_{201-252}$ and EZH2 with individual mutations were immobilized on the His-Pur Cobalt resin and allowed to interact overnight with GST-VAV$_{1-172}$ or GST alone (control) in PBS buffer pH 7.4. Following this, the beads were extensively washed and subject to SDS-PAGE and Western blot analysis using anti-His (h-3, Santa Cruz) and anti-GST (B-14, Santa Cruz) antibodies from Santa Cruz to analyze the EZH2-VAV complex. To determine the interaction between full-length EZH2 and VAV, VAV1 protein was immunoprecipitated from the cytosolic fraction of Jurkat cells with α-VAV1 antibody from Santa Cruz (C-14) and immobilized on Protein G Sepharose. Similarly, Flag-tagged EZH2 WT or EZH2 with individual mutations were immunoprecipitated from HEK-293 cells with α-Flag antibody from Sigma (M2) and Protein G Sepharose beads. They were then eluted from the beads by using a competing FLAG peptide. Equal quantities of the WT EZH2 or EZH2 with individual mutations were incubated overnight with VAV1 in a BC-100 buffer. The beads were then extensively washed and analysed by Western blotting with α-EZH2 antibodies from Santa Cruz (N-20) and VAV-specific antibodies.

Immunoblotting and antibodies

Immunoblotting was performed according to standard protocols with the following antibodies; EZH2 (D2C9, Cell Signaling), EED (AA19, Millipore), SUZ12 (Ab12073, Abcam), CDYL (ab5188, Abcam), Actin (C-11, Santa Cruz), VAV2 (EP1067Y, Abcam), VAV pY174 (sc-16408-R, Santa Cruz), H3 (ab1791, Abcam), H3k27me3 (07-449, Abcam), STAT3 (F2, Santa

Cruz), STAT3 pY705 (D3A7, Cell Signaling), STAT3 pS727 (ab30647), EGFR (sc-03-G, Santa Cruz), EGFR pY1173 (sc-12351-R, Santa Cruz), AKT1/2 (N-19, Santa Cruz), AKT pS473 (193H12, Cell Signaling), Tubulin (AA13, Santa Cruz), LaminB (C-20, Santa Cruz), Talin1, N-terminal (TA205, Millipore), Talin1 (also used for IP, 8d4, Santa Cruz), Talin1 pS425 (5426S, Cell Signaling), FAK (2601P, Cell Signaling), FAK pY397 (3283S, Cell Signaling), FAK pY925 (3284S, Cell Signaling), FAK pY576/577 (3281S, Cell Signaling), GAPDH (6C5, Ambion), pan-methyl-lysine (ab3715, Abcam), Src (32G6, Cell Signaling), Src pY416 (D49G4, Cell Signaling), Src pY527 (2105P, Cell Signaling), ERK (137F5, Cell Signaling), ERK1/2 pT202/Y204 (D13.14.4E, Cell Signaling), Paxillin pY-118 (2541S, Cell Signaling), PAK1/2/3 (2604P, Cell Signaling), PAK1 pS144 (2606P, Cell Signaling), FACS antibodies; human CD44 (IM7, Biolegend), CD24 (M1/69, Biolegend)

Surface plasmon resonance

Quantitative analysis of the EZH2-VAV interaction using $VAV1_{1-172}$ and mutant forms of $EZH2_{201-252}$ was performed using BIACORE 3000. A 4000 RU anti-GST surface was prepared on a standard CM5 sensor chip (GE Healthcare) by amine-coupling anti-GST antibody (GST capture kit, GE Healthcare) on all four flow cells. Subsequent interaction assays were performed by capturing approximately 2700 RU of GST-$VAV1_{1-172}$ (equivalent GST on separate flow cell as control) using a buffer of 10 mM PBS, 137 mM NaCl, 2.7 mM KCl, pH 7.4 at 25^{o}C. A typical cycle of interaction comprises: (1) stabilization in running buffer, (2) capturing of ligand, (3) stabilization after ligand capture, (4) interaction with analytes, (5) dissociation of analytes, and (6) regeneration of surface with 1:10 phosphoric acid. Several concentrations (32 nM – 16.7 µM) of EZH2-HIS and its mutants were injected across control and GST-VAV1-captured surface for 2 min at a flow rate of 30 µl/min. The equilibrium response for each EZH2-HIS concentration was converted to surface ligand occupancy by

dividing R_{max}. These occupancy responses were plotted against the log concentration of EZH2-HIS as standard dose-response curves and a steady-state model was used to fit the points. The affinity constants were interpolated from 50% occupancy.

In vitro methyltransferase assay

Full-length wild-type EZH2 or EZH2 with the individual mutations were incubated with recombinant histone octamer (0.2 µg) using S-adenosyl-I-[methyl-^{3}H] methionine ([^{3}H]-SAM) or cold SAM as the methyl donor. The reaction was carried out in methylation buffer (50 mM Tris, pH8.0) containing 0.5 mM DTT, 7.6 µM (15 µci) of SAM at 30 ^{o}C for 1 h. The reaction was stopped by addition of 2x SDS gel loading buffer followed by SDS-PAGE. The gel was either stained with Coomassie Blue Brilliant and enhanced by EN3HANCE (Perkin-Elmer Life Sciences) for X-ray film exposure or transferred to a PVDF membrane (Millipore) and probed with anti-H3K27me3 or anti-pan-methyl-lysine antibodies.

Immunofluorescence staining

Immunofluorescence staining was performed according to standard protocols. H16N2 cells stably expressing different forms of EZH2, as described, were plated on glass slides at a concentration of 3.0 x 10^5 per well. Cells were starved of EGF for 3 days and then stimulated with EGF (30 ng/ml) for 12 h. They were then fixed with 2% paraformaldehyde for 30 min and permeabilized with 0.1% Triton X-100 for 2 min, washed and stained for F-actin with fluorescein isothiocyanate- or rhodamine red-conjugated or Alexa Fluor® 647 Phalloidin (molecular Probes). Slides were blocked with 10% normal donkey serum for 1 hour, then stained with anti-Paxillin (05-417PC12-305, Millipore) or anti-GFP (ab290, Abcam) antibodies in PBS/BSA at 4^{o}C overnight and then incubated with anti-mouse Rhodamine secondary antibody or anti-rabbit FITC (Jackson ImmunoResearch) antibody in PBS/BSA at room temperature for 1 hour. Slides were then mounted with DAPI (Invitrogen) containing

mounting medium (Ibidi GmbH) and observed using a 60× objective by immunofluorescence microscopy.

Live cell imaging for adhesion assembly and disassembly

Live cell imaging was performed using an inverted epifluorescent microscope Eclipse Ti microscope (Nikon) at 100× magnification. Images of live cells were recorded at a time-lapse interval of 2 min over a period of 1-2 h using an ORCA-Flash 4.0 camera (Hamatsu) with the MetaMorph Imaging System 7.8 (Molecular Devices). Time-lapse images were reconstructed as image stacks (MetaMorph). The background and photobleaching-corrected integrated fluorescence intensities in individual adhesion based on mCherry-Paxillin or GFP-Talin1 were determined manually using the "integrated density" measurement function of ImageJ. Focal complexes and focal adhesions with a size between 0.2 μm^2 and 19 μm^2 were scored. Integrated fluorescence intensities were plotted against time. A linear trendline fitted to the data was used to identify the period of linear assembly or disassembly of the adhesions. The rate constants of assembly (k_a) and disassembly (k_d) were calculated using the equations: $k_a = -\frac{ln(I/I_0)}{(t-t_0)}$ and $k_d = -\frac{ln(I_0/I)}{(t_0-t)}$, where *[I]* is the intensity of the adhesion at a given time point and *[I_0]* is the initial intensity of the adhesion. Rate constants were used to calculate half-life ($t_{1/2}$) values for adhesion assembly and disassembly using the equation: $t_{1/2} = \frac{\ln(2)}{k_a \text{ or } k_d}$

Anchorage–independent cell-growth assay

Anchorage-independent growth of H16N2 cells stably expressing various forms of EZH2 was tested in 24-well culture plates. Two hundred μl of F12 culture medium with 0.6% soft-agar was first plated into each well of a 24-well plate. After the agar solidified, each well received another 1 ml of 0.4% agar in culture medium containing 15,000 cells. After 25-30 days,

colonies were fixed and stained with 0.1% crystal violet. The number of colonies from triplicate wells for each cell line was determined using Image J software.

Mammary tumor models

Inguinal mammary fat pads of 6-week-old female NOD-SCID mice or BALB/c mice were inoculated with 7 x 10^6 of MDA-MB-231 cells or 1 x 10^6 of EZH2 knockdown 4T1 cells expressing different forms of EZH2 variants. Tumor growth was monitored at the indicated time points by measuring tumor dimensions (length and width) using Vernier calipers or fluorescence intensity using IVIS Spectrum CT (PerkinElmer) 24 hrs after i.v. injection of 2-DG750 (10 nmoles/100 µl). To estimate the tumor volume measured by Vernier calipers, we assumed the tumor shape was a cylinder and calculated the volume of the cylinder by multiplying the ellipse area by the estimated depth (we took whichever was larger between width and length as depth). The actual depth and tumor volume cannot be measured in live animals using this method. At the end of the experimental period, mice were sacrificed, tumors and other organs were resected, and either weighed and fixed with 10% paraformaldehyde for tissue sectioning or the fluorescence intensity was measured by IVIS Spectrum CT. All mice were bred and maintained under specific pathogen-free conditions at the animal facility of the School of Biological Sciences, Nanyang Technological University. All mouse protocols were conducted in accordance with the guidelines of the NTU Institutional Animal Care and Use Committee (IACUC).

RNA-seq

EZH2 knockdown H16N2 cells transduced with control (MIG) or retrovirus to express the indicated EZH2 variants were seeded in 6-well plates for 48 hours. RNA from these cells was isolated using the QIAshredder (Qiagen) and RNeasy (Qiagen) kits. The total RNA quality and

RIN value was confirmed with a 2100 Bioanalyzer (Agilent) using RNA 6000 Nano chip (Agilent). Library preparation was done using the Illumina® TruSeq Stranded mRNA LT Sample Prep Kit, by following the manufacturer's protocol. Each sample was linearized, denatured and loaded into 2 lanes of Single-Read flowcell using the Illumina cBOT. The cDNAs were attached to the flowcell surfaces and amplified to clusters and attached with sequencing primers. The flowcell was then loaded into the Illumina HiSeq2000 Sequencer with the Illumina® HiSeq SBS Kit and run at 1x101 cycles, generating Single-Read 100 base-pair reads. The images were captured by the HiSeq Control Software (HCS), and the Real Time Analysis (RTA) software converted the images into Cycle Intensity Files (CIF) and, subsequently, Basecall (bcl) files. RNA-Seq data were mapped against the human genome version hg19 with TopHat2, using the GENCODE Release 19 version of gene annotations. FPKM values were extracted for the RNA-Seq analysis. Reads were counted using the R package GenomicAlignments (mode='Union', inter.feature=FALSE), only primary read alignments were retained. Differentially expressed genes were identified using DESeq2. The differentially expressed genes for the 4 conditions (MIG, EZH2ΔNLS, EZH2H689A, and EZH2-VAV/MT) were calculated in comparison to EZH2. If the log2-fold change of a gene was higher than 0 and its adjusted p-value was smaller than 0.05, the gene was highlighted as upregulated. If the log2-fold change of a gene was smaller than 0 and its adjusted p-value was smaller than 0.05, the gene was highlighted as downregulated. The heatmap was created by using ggplot2_1.0.0 (for detailed references, see Supplemental Experimental Procedures). Normalized FPKM values of these significantly expressed genes were shown as a heatmap in Fig. 6A. The data were normalized such that the sum of the expression estimates of each gene was 1 (relative expression) and the normalized FPKM value from the EZH2 sample was subtracted from the other samples in order to show the EZH2 sample as the baseline. Functional ontology analysis of differentially expressed genes (more than 0.3 of log2-fold increase or

decrease and $p<0.05$ relative to results obtained in wild-type EZH2-expressing cells) in cytosolic EZH2-expressing cells were classified using the DAVID Bioinformatics Resources online functional annotation tool (http://david.abcc.ncifcrf.gov/summary.jsp), then grouped by functional categories (see corresponding Fig. 6 source data file).

Accession codes:
ArrayExpress website
http://www.ebi.ac.uk/arrayexpress/help/how_to_search.html#Login
Username: Reviewer_E-MTAB-3599
Password: 3svJrgcv

Real-time PCR

RNA was extracted from H16N2 cells, as above, and subsequently converted to complementary DNA (cDNA) using the SuperScript III First-Strand Synthesis Kit (Invitrogen) with random hexamers as primers. The cDNA abundance was analyzed by SYBR green real-time PCR with respective primers using the StepOnePlus™ Real-Time PCR System (Applied Biosystems). GAPDH was used as a reference gene. Fold change was calculated using the formula $2^{-\Delta\Delta Ct}$, where $\Delta\Delta Ct=\Delta Ct$ sample – ΔCt reference sample and $\Delta Ct=Ct$ target gene – Ct reference gene. The following primer sequences were used for each gene:

STAT1: 5’ CAGCTTGACTCAAAATTCCTGGA 3’, 5’ TGAAGATTACGCTTGCTTTTCCT 3’
STAT2: 5’ CCAGCTTTACTCGCACAGC 3’, 5’ AGCCTTGGAATCATCACTCCC 3’
STAT3: 5’ ACCAGCAGTATAGCCGCTTC 3’, 5’ GCCACAATCCGGGCAATCT 3’
IRF1: 5’ ATGCCCATCACTCGGATGC 3’, 5’ CCCTGCTTTGTATCGGCCTG 3’
c-Fos: 5’ GGGGCAAGGTGGAACAGTTAT 3’, 5’ CCGCTTGGAGTGTATCAGTCA 3’
MMP7: 5’ GAGTGAGCTACAGTGGGAACA 3’, 5’ CTATGACGCGGGAGTTTAACAT 3’
MMP13: 5’ TCCTGATGTGGGTGAATACAATG 3’, 5’ GCCATCGTGAAGTCTGGTAAAAT 3’
TMPRSS15: 5’ GCCAGTGGATCATACGTGTAAA 3’, 5’ GTGCCAGGATTAGTTTCCCAAAT 3’
IFI27: 5’ GTGGACGTTCGGGAGCTAC 3’, 5’ ACTGGCCGATTTGGCACAG 3’

AUTHOR CONTRIBUTIONS

N.V. identified EZH2-mediated Talin methylation sites, initiated the project and established essential collaborations; N.V. and J.F.W. designed and conducted most of the experiments, including the *in vitro* methyltransferase assays, identification of the EZH2-VAV interaction mutant and performed experiments in mammary epithelial cells, in breast cancer animal models and were also involved in the data interpretation and manuscript preparation; K.P.T. and M.S.M. did computational modeling and predicted residues critical for EZH2–VAV interaction; H.H.C and V.C.L.L. were involved in experiments using breast cancer animal models; E.C. and J.G. performed differential expression analysis of RNA-seq data; Y.H.Y. and S.G.S. performed the SPR analysis A.A. and L.N. produced recombinant octamers and I.S. designed and interpreted the experiments and wrote the manuscript.

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

We thank S. Strömblad for critical reading, discussions and comments on the manuscript; Y. Zhang (Harvard University) for the EZH2, SUZ12 and EED baculovirus expression vectors; we thank Y. Or for providing BALB/c mice and technical support, L.Y. Lu for advice on statistical analysis and A. Sullivan from Obrizus Communications for editing the manuscript. This work was supported by the Singapore Agency for Science Technology and Research (A*STAR) through a Biomedical Research Council (BMRC) grant (10/1/22/19/666) to L.N., Singapore Ministry of Education (MOE2009-T2-1-034 and MOE2013-T2-2-038) and Ministry of Health, National Medical Research Council (NMRC-CBRG/0057/201) to I.S.

"Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)".

REFERENCES

1 Aghazadeh B, Lowry WE, Huang XY, Rosen MK (2000). Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation. *Cell* **102:** 625-633.

2 Boddicker RL, Razidlo GL, Dasari S, Zeng Y, Hu G, Knudson RA *et al* (2016). Integrated mate-pair and RNA sequencing identifies novel, targetable gene fusions in peripheral T-cell lymphoma. *Blood* **128:** 1234-1245.

3 Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS *et al* (2016). Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* **48:** 607-616.

4 Cao Q, Yu J, Dhanasekaran SM, Kim JH, Mani RS, Tomlins SA *et al* (2008). Repression of E-cadherin by the polycomb group protein EZH2 in cancer. *Oncogene* **27:** 7274-7284.

5 Cao Q, Wang X, Zhao M, Yang R, Malik R, Qiao Y *et al* (2014). The central role of EED in the orchestration of polycomb group complexes. *Nat Commun* **5:** 3127.

6 Chang CJ, Yang JY, Xia W, Chen CT, Xie X, Chao CH *et al* (2011). EZH2 promotes expansion of breast tumor initiating cells through activation of RAF1-beta-catenin signaling. *Cancer Cell* **19:** 86-100.

7 Chang KH, Sanchez-Aguilera A, Shen S, Sengupta A, Madhu MN, Ficker AM *et al* (2012). Vav3 collaborates with p190-BCR-ABL in lymphoid progenitor leukemogenesis, proliferation, and survival. *Blood* **120:** 800-811.

8 Chen YH, Hung MC, Li LY (2012). EZH2: a pivotal regulator in controlling cell differentiation. *Am J Transl Res* **4:** 364-375.

9 Chen YT, Zhu F, Lin WR, Ying RB, Yang YP, Zeng LH (2016). The novel EZH2 inhibitor, GSK126, suppresses cell migration and angiogenesis via down-regulating VEGF-A. *Cancer Chemother Pharmacol* **77:** 757-765.

10 Chou RH, Yu YL, Hung MC (2011). The roles of EZH2 in cell lineage commitment. *Am J Transl Res* **3:** 243-250.

11 Chung J, Uchida E, Grammer TC, Blenis J (1997). STAT3 serine phosphorylation by ERK-dependent and -independent pathways negatively modulates its tyrosine phosphorylation. *Mol Cell Biol* **17:** 6508-6516.

12 Citterio C, Menacho-Marquez M, Garcia-Escudero R, Larive RM, Barreiro O, Sanchez-Madrid F *et al* (2012). The rho exchange factors vav2 and vav3 control a lung metastasis-specific transcriptional program in breast cancer cells. *Sci Signal* **5:** ra71.

13 Dasgupta M, Dermawan JK, Willard B, Stark GR (2015). STAT3-driven transcription depends upon the dimethylation of K49 by EZH2. *Proc Natl Acad Sci U S A* **112:** 3985-3990.

14 De Craene B, Berx G (2013). Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer* **13:** 97-110.

15 de Vries NA, Hulsman D, Akhtar W, de Jong J, Miles DC, Blom M *et al* (2015). Prolonged Ezh2 Depletion in Glioblastoma Causes a Robust Switch in Cell Fate Resulting in Tumor Progression. *Cell Rep* **10:** 383-397.

16 Ezhkova E, Pasolli HA, Parker JS, Stokes N, Su IH, Hannon G *et al* (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell* **136:** 1122-1135.

17 Gialeli C, Theocharis AD, Karamanos NK (2011). Roles of matrix metalloproteinases in cancer progression and their pharmacological targeting. *FEBS J* **278:** 16-27.

18 Gonzalez ME, Moore HM, Li X, Toy KA, Huang W, Sabel MS *et al* (2014). EZH2 expands breast stem cells through activation of NOTCH1 signaling. *Proc Natl Acad Sci U S A* **111:** 3098-3103.

19 Groysman M, Nagano M, Shaanan B, Katzav S (1998). Mutagenic analysis of Vav reveals that an intact SH3 domain is required for transformation. *Oncogene* **17:** 1597-1606.

20 Gunawan M, Venkatesan N, Loh JT, Wong JF, Berger H, Neo WH *et al* (2015). The methyltransferase Ezh2 controls cell adhesion and migration through direct methylation of the extranuclear regulatory protein talin. *Nat Immunol* **16:** 505-516.

21 Harrington EA, Fanidi A, Evan GI (1994). Oncogenes and cell death. *Curr Opin Genet Dev* **4:** 120-129.

22 He A, Shen X, Ma Q, Cao J, von Gise A, Zhou P *et al* (2012). PRC2 directly methylates GATA4 and represses its transcriptional activity. *Genes Dev* **26:** 37-42.

23 Herranz N, Pasini D, Diaz VM, Franci C, Gutierrez A, Dave N *et al* (2008). Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. *Mol Cell Biol* **28:** 4772-4781.

24 Hobert O, Jallal B, Ullrich A (1996). Interaction of Vav with ENX-1, a putative transcriptional regulator of homeobox gene expression. *Mol Cell Biol* **16:** 3066-3073.

25 Hobert O, Sures I, Ciossek T, Fuchs M, Ullrich A (1996). Isolation and developmental expression analysis of Enx-1, a novel mouse Polycomb group gene. *Mech Dev* **55:** 171-184.

26 Hock H (2012). A complex Polycomb issue: the two faces of EZH2 in cancer. *Genes Dev* **26:** 751-755.

27 Hsieh FC, Cheng G, Lin J (2005). Evaluation of potential Stat3-regulated genes in human breast cancer. *Biochem Biophys Res Commun* **335:** 292-299.

28 Huq MD, Tsai NP, Khan SA, Wei LN (2007). Lysine trimethylation of retinoic acid receptor-alpha: a novel means to regulate receptor function. *Mol Cell Proteomics* **6:** 677-688.

29 Jain N, Iyer KV, Kumar A, Shivashankar GV (2013). Cell geometric constraints induce modular gene-expression patterns via redistribution of HDAC3 regulated by actomyosin contractility. *Proc Natl Acad Sci U S A* **110:** 11349-11354.

30 Jiao L, Liu X (2015). Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* **350:** aac4383.

31 Juan AH, Kumar RM, Marx JG, Young RA, Sartorelli V (2009). Mir-214-dependent regulation of the polycomb protein Ezh2 in skeletal muscle and embryonic stem cells. *Mol Cell* **36:** 61-74.

32 Kamminga LM, Bystrykh LV, de Boer A, Houwer S, Douma J, Weersing E *et al* (2006). The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107:** 2170-2179.

33 Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J *et al* (2015). Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat Genet* **47:** 1304-1315.

34 Katzav S, Martin-Zanca D, Barbacid M (1989). vav, a novel human oncogene derived from a locus ubiquitously expressed in hematopoietic cells. *Embo J* **8:** 2283-2290.

35 Kim E, Kim M, Woo DH, Shin Y, Shin J, Chang N *et al* (2013). Phosphorylation of EZH2 activates STAT3 signaling via STAT3 methylation and promotes tumorigenicity of glioblastoma stem-like cells. *Cancer Cell* **23:** 839-852.

36 Kim KH, Roberts CW (2016). Targeting EZH2 in cancer. *Nat Med* **22:** 128-134.

37 Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA *et al* (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc Natl Acad Sci U S A* **100:** 11606-11611.

38 Knutson SK, Wigle TJ, Warholic NM, Sneeringer CJ, Allain CJ, Klaus CR *et al* (2012). A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat Chem Biol* **8:** 890-896.

39 Lee JM, Lee JS, Kim H, Kim K, Park H, Kim JY *et al* (2012). EZH2 generates a methyl degron that is recognized by the DCAF1/DDB1/CUL4 E3 ubiquitin ligase complex. *Mol Cell* **48:** 572-586.

40 Lee ST, Li Z, Wu Z, Aau M, Guan P, Karuturi RK *et al* (2011). Context-specific regulation of NF-kappaB target gene expression by EZH2 in breast cancers. *Mol Cell* **43:** 798-810.

41 Lee TC, Ziff EB (1999). Mxi1 is a repressor of the c-Myc promoter and reverses activation by USF. *J Biol Chem* **274:** 595-606.

42 Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM *et al* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125:** 301-313.

43 Lock JG, Mamaghani MJ, Shafqat-Abbasi H, Gong X, Tyrcha J, Stromblad S (2014). Plasticity in the macromolecular-scale causal networks of cell migration. *PLoS One* **9:** e90593.

44 Matsuguchi T, Inhorn RC, Carlesso N, Xu G, Druker B, Griffin JD (1995). Tyrosine phosphorylation of p95Vav in myeloid cells is regulated by GM-CSF, IL-3 and steel factor and is constitutively increased by p210BCR/ABL. *EMBO J* **14:** 257-265.

45 McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS *et al* (2012). EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492:** 108-112.

46 Menacho-Marquez M, Garcia-Escudero R, Ojeda V, Abad A, Delgado P, Costa C *et al* (2013). The Rho exchange factors Vav2 and Vav3 favor skin tumor initiation and promotion by engaging extracellular signaling loops. *PLoS Biol* **11:** e1001615.

47 Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R *et al* (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet* **42:** 181-185.

48 Nagano M, Hoshino D, Koshikawa N, Akizawa T, Seiki M (2012). Turnover of Focal Adhesions and Cancer Cell Migration. *International Journal of Cell Biology* **2012**.

49 Natsume A, Ito M, Katsushima K, Ohka F, Hatanaka A, Shinjo K *et al* (2013). Chromatin regulator PRC2 is a key regulator of epigenetic plasticity in glioblastoma. *Cancer Res* **73:** 4559-4570.

50 Nicolaidou V, Wong MM, Redpath AN, Ersek A, Baban DF, Williams LM *et al* (2012). Monocytes induce STAT3 activation in human mesenchymal stem cells to promote osteoblast formation. *PLoS One* **7:** e39871.

51 Normanno N, De Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR *et al* (2006). Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* **366:** 2-16.

52 Ntziachristos P, Tsirigos A, Van Vlierberghe P, Nedjic J, Trimarchi T, Flaherty MS *et al* (2012). Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nat Med* **18:** 298-301.

53 Pasini D, Bracken AP, Jensen MR, Denchi EL, Helin K (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J* **23:** 4061-4071.

54 Rangel LB, Agarwal R, D'Souza T, Pizer ES, Alo PL, Lancaster WD *et al* (2003). Tight junction proteins claudin-3 and claudin-4 are frequently overexpressed in ovarian cancer but not in ovarian cystadenomas. *Clin Cancer Res* **9:** 2567-2575.

55 Ren G, Baritaki S, Marathe H, Feng J, Park S, Beach S *et al* (2012). Polycomb protein EZH2 regulates tumor invasion via the transcriptional repression of the metastasis suppressor RKIP in breast and prostate cancer. *Cancer Res* **72:** 3091-3104.

56 Rodriguez JL, Sandoval J, Serviddio G, Sastre J, Morante M, Perrelli MG *et al* (2006). Id2 leaves the chromatin of the E2F4-p130-controlled c-myc promoter during hepatocyte priming for liver regeneration. *Biochem J* **398:** 431-437.

57 Rudolph MG, Stanfield RL, Wilson IA (2006). How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* **24:** 419-466.

58 Sauvageau M, Sauvageau G (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* **7:** 299-313.

59 Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT *et al* (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol* **4:** 124-131.

60 Su IH, Dobenecker MW, Dickinson E, Oser M, Basavaraj A, Marqueron R *et al* (2005). Polycomb group protein ezh2 controls actin polymerization and cell signaling. *Cell* **121:** 425-436.

61 Su IH, Tarakhovsky A (2006). Lysine methylation and 'signaling memory'. *Curr Opin Immunol* **18:** 152-157.

62 Tan J, Yang X, Zhuang L, Jiang X, Chen W, Lee PL *et al* (2007). Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev* **21:** 1050-1063.

63 Tanos B, Rodriguez-Boulan E (2008). The epithelial polarity program: machineries involved and their hijacking by cancer. *Oncogene* **27:** 6939-6957.

64 Thalappilly S, Soubeyran P, Iovanna JL, Dusetti NJ (2010). VAV2 regulates epidermal growth factor receptor endocytosis and degradation. *Oncogene* **29:** 2528-2539.

65 Toh TB, Lim JJ, Chow EKH (2017). Epigenetics in cancer stem cells. *Mol Cancer* **16**.

66 Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG *et al* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419:** 624-629.

67 Vergani L, Grattarola M, Nicolini C (2004). Modifications of chromatin structure and gene expression following induced alterations of cellular shape. *Int J Biochem Cell Biol* **36:** 1447-1461.

68 Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C *et al* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439:** 871-874.

69 Wassef M, Rodilla V, Teissandier A, Zeitouni B, Gruel N, Sadacca B *et al* (2015). Impaired PRC2 activity promotes transcriptional instability and favors breast tumorigenesis. *Genes Dev* **29:** 2547-2562.

70 Webb DJ, Donais K, Whitmore LA, Thomas SM, Turner CE, Parsons JT *et al* (2004). FAK-Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly. *Nat Cell Biol* **6:** 154-161.

71 Wierstra I, Alves J (2008). The c-myc promoter: still MysterY and challenge. *Adv Cancer Res* **99:** 113-333.

72 Xu K, Wu ZJ, Groner AC, He HH, Cai C, Lis RT *et al* (2012). EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science* **338:** 1465-1469.

73 Yang J, Huang J, Dasgupta M, Sears N, Miyagi M, Wang B *et al* (2010). Reversible methylation of promoter-bound STAT3 by histone-modifying enzymes. *Proc Natl Acad Sci U S A* **107:** 21499-21504.

74 Yoo KH, Hennighausen L (2012). EZH2 methyltransferase and H3K27 methylation in breast cancer. *Int J Biol Sci* **8:** 59-65.

75 Yu H, Jove R (2004). The STATs of cancer--new molecular targets come of age. *Nat Rev Cancer* **4:** 97-105.

76 Yu J, Yu J, Rhodes DR, Tomlins SA, Cao X, Chen G *et al* (2007). A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. *Cancer Res* **67:** 10657-10663.

77 Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D *et al* (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481:** 157-163.

78 Zhang Y, Yang X, Gui B, Xie G, Zhang D, Shang Y *et al* (2011). Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27. *J Biol Chem* **286:** 42414-42425.

79 Zhao E, Maj T, Kryczek I, Li W, Wu K, Zhao L *et al* (2016). Cancer mediates effector T cell dysfunction by targeting microRNAs and EZH2 via glycolysis restriction. *Nat Immunol* **17:** 95-103.

80 Zhu P, Wang Y, Huang G, Ye B, Liu B, Wu J *et al* (2016). lnc-beta-Catm elicits EZH2-dependent beta-catenin stabilization and sustains liver CSC self-renewal. *Nat Struct Mol Biol* **23:** 631-639.

FIGURE LEGENDS

Figure 1 Structural prediction and identification of the critical residues on EZH2 for EZH2-VAV interaction

(**a**) Purified GST-tagged $VAV1_{1-172}$ and 6xHis tagged $EZH2_{201-252}$ were immunoprecipitated with either His-Pur Cobalt resin or glutathione beads. Immunoblots with the indicated antibodies are shown. 'C' indicates GST control incubated with EZH2 fragment and 'EV' indicates EZH2 fragment incubated with VAV1 fragment. (**b-d**) EZH2-VAV1 interaction was characterized by surface plasmon resonance (SPR) analysis. (**b**) Purified EZH2-$HIS_{201-252}$ and GST-$VAV1_{1-172}$ were analyzed by Coomassie-stained gel before being subjected to SPR analysis. (**c**) $VAV1_{1-172}$ interactions with $EZH2_{201-252}$ as determined by SPR. Sensograms obtained were buffer and reference flow cell corrected. (**d**) Equilibrium responses were plotted against the log of EZH2-$HIS_{201-252}$ concentration. The estimated affinity constant of the interaction between EZH2-$HIS_{201-252}$ and GST-$VAV1_{1-172}$ was 600 nM based on the steady-state model. (**e**) Comparative model of the VAV1-EZH2 interaction. VAV1 is shown in surface representation while the interacting helix of EZH2 is shown as ribbon. The surface is colored by its electrostatic potential. Positive and negative regions are shaded blue and red, respectively. Detailed information for the structural analysis is listed in the Supplementary Methods. (**f**) Predicted mutations in EZH2 and the complementary mutations in VAV1 to disrupt EZH2-VAV1 interaction. (**g**) Sequence alignment of the EZH2-interacting stretch of human VAV1, VAV2, and VAV3. Alignment positions marked with an '*' indicate sequence conservation, while '**:**' and '**.**' denote levels of sequence similarity (details see Supplementary Fig. 1b). (**h**) The interactions of HIS-tagged $EZH2_{201-252}$ point mutants with wild-type $VAV1_{1-172}$ fragment were determined by immobilizing EZH2 using His-Pur Cobalt resin and incubating with purified VAV1 fragments. Bar graph on the right summarizes the fold changes in VAV1 co-precipitated by EZH2 mutants relative to the amount pulled down by wild-type EZH2 from 3 independent experiments. Means ± standard deviation (s.d.) are shown. $^{**}P$ = 0.001, 0.007, 0.0004, 0.002, 0.0002 (left to right), $^{*}P$ = 0.03 (two-tailed Student's t test). (**i**) $VAV1_{1-172}$ interactions with $EZH2_{201-252}$ mutants as determined by SPR. Interactions were performed as in **c**. (**j**) Empty vector (EV), wild-type EZH2, and EZH2-VAV/MT fused with C-terminal GFP-FLAG-tags were transiently expressed in BOSC23 cells and EZH2 fusion proteins were purified using anti-FLAG antibody under stringent conditions free of interacting proteins (see input controls). The amounts of purified EZH2 variants were adjusted and incubated with BOSC23 lysates to interact with VAV2. The co-precipitated VAV2 was determined by

specific antibody. Bar graph summarizes the fold changes in VAV2 (Co-p VAV2) co-precipitated by EZH2-VAV/MT relative to that pulled down by wild-type EZH2 from 3 independent experiments. Mean ± standard deviation (s.d.) for each condition is shown. ** P = 0.005 (two-tailed Student's t test). (**k**) EZH2 was immunoprecipitated from total cell extracts of EZH2 knockdown MCF10A cells reconstituted with the indicated EZH2 variants and used in an *in vitro* methyltransferase assay with cold S-adenosyl methionine (SAM) as a methyl donor and recombinant histone octamer as a substrate. H3K27me3 was determined by immunoblot with specific antibody. (**l**) Histone methyltransferase activities of wild-type EZH2 and the VAV targeting mutant EZH2-K234/245A (EZH2-VAV/MT) immunoprecipitated from reconstituted EZH2 knockdown MCF10A cells were measured as in **k**, except with titrated amounts of EZH2, as indicated. Immunoblots shown in this Figure are representative of more than 3 independent experiments.

Figure 2 EZH2 lacking its NLS shows elevated oncogenic capacity

(**a**) Wild-type un-transduced (Un), endogenous EZH2 knockdown (Endo-KD), VAV2 knockdown (VAV2-KD), or H16N2 cells transduced with control (MIG) or retrovirus to express the indicated EZH2 variants were tested for anchorage independent cell growth. Data represent means ± s.d., (3 independent experiments with technical triplicates). Two-tailed Student's t-test with equal variance, Left graph: ** P = 0.0005, 0.002, 0.001, *** $P = 4.8 \times 10^{-6}$, 8.5×10^{-6} (left to right). Middle graph: *** $P = 1.2 \times 10^{-5}$. Right graph: *** $P = 2 \times 10^{-13}$. (**b-d**) MDA-MB-231 (MDA) human breast cancer cells transduced with EZH2 variants were inoculated into the mammary fat pads of 6-week old female NOD-SCID mice. (**b**) The tumor volumes at different time points were estimated based on the width and length measurements using a Vernier caliper, as described in Methods. Statistical analyses were performed between wild-type EZH2 or EZH2ΔNLS and EZH2-VAV/MT or MIG controls. (**c**) Tumor weight was measured after sacrificing the animal at the end of experiment. Representative images of resected primary tumors (**c**, right panel) and draining lymph nodes (**d**, top panel) are shown. Data represent means ± standard error of the mean (s.e.m.). One-tailed Student's t-test with equal variance, *P=0.02, 0.04, **P=0.002, 0.002, left to right in **b**, ** P= 0.0003, 0.0002, 0.006, 0.002, top to bottom in **c**, * P = 0.02 in **d**. n = 6 mice per group for experiments shown in **b-d**. (**e-g**) Mouse 4T1 breast tumor cells, in which endogenous EZH2 was knocked down and subsequently transduced with retrovirus for the expression of wild-type EZH2 or the indicated EZH2 variants, were inoculated into the mammary fat pads of 6-week old female

BALB/c mice. Tumor growth was monitored by manual measurement and visualized by IVIS Spectrum CT using 2-DG probe on the days indicated. (**e**) Relative 2-DG fluorescence signals of the primary tumors and metastases in the thoracic region of the animals at 37 days after inoculation are shown in bar graphs. Data represent means ± s.e.m. Two-tailed Student's t-test with equal variance, ** $P = 0.002$, * $P = 0.048$. (**f**) Representative images of fluorescence labelled tumors in live animals at the indicated time points are shown. (**g**) At day 37, mice were sacrificed and the internal organs were imaged. Top and bottom panels represent data from two independent sets of experiments. The modest tumor growth enhancing effect of EZH2H689A in **f** and **g**, could be due to residual enzymatic activity (20-30%) combined with high expression levels. "#" indicates neck lymph node metastases found in all mice that received EZH2ΔNLS expressing 4T1 cells. Representative images from one of 2 independent experiments are shown. $n = 6$ animals per group for experiments shown in **e-g**.

Figure 3 EZH2 promotes adhesion turnover in a VAV interaction dependent manner

(**a**) H16N2 cells were transduced with control virus (MIG) or retroviruses for the expression of the indicated EZH2 variants and subsequently endogenous EZH2 was knocked down by shRNA. "Un" indicates un-transduced wild-type control cells. Cells were prepared as specified in **Supplementary Fig. 4**. Adhesion structures were visualized using a specific antibody against Paxillin (red). Phalloidin and DAPI were used to stain F-actin (green) and nuclei (blue), respectively. Scale bars, 20 μm. Representative images from one of 3 independent experiments are shown. Quantification of various types of adhesion structures in these cells is summarized in **Supplementary Figure 4**. (**b**) The empirical cumulative distribution function curve (CDF) shows the direction of the changes in the size of adhesion structures in H16N2 cells expressing EZH2 variants as in **a**. The adhesion structures were determined by Paxillin staining. Kolmogorov-Smirnov tests were performed between groups using software available online (http://www.wessa.net/rwasp_Reddy-Moores%20K-S%20Test.wasp). The P values are all <0.00001. The graph in the lower panel shows cumulative (cumul.) numbers of adhesion structures in each group. Representative results from one of the 3 independent experiments with 60 cells per group are shown. (**c**) H16N2 cells expressing EZH2 variants, as in **a**, were transfected with plasmids to express mCherry-Paxillin and plated on fibronectin coated glass slides, then analysed by time-lapse live imaging. Half-life values of adhesion assembly and disassembly were calculated based

on mCherry-Paxillin signal, as described[70]. Data represent means ± s.d. of 15 adhesions on 4-5 cells per group. Two-tailed Student's t-test with equal variance, ** P = 0.007, *** P = 6.4 x 10^{-9}, 1.4 x 10^{-7}, 4.4 x 10^{-5} (top to bottom). (**d**) H16N2 cells were left un-transfected or co-transfected with EZH2 shRNA and empty vector (EV) or vectors to express variants of EZH2-GFP protein. Cell spreading of GFP positive cells, compared to cells transduced with EV, was quantified. At least 3 experiments and more than 20 cells were scored for each group in each experiment. Data represent means ± s.e.m. Two-tailed Student's t-test with *equal variance, ** P = 0.0006, 0.0007, 0.002, 0.0005 (left to right), * P = 0.03.*

Figure 4 EZH2-mediated Talin1 methylation regulates adhesion dynamics in mammary epithelial cells (**a**) Cytosolic extracts of EZH2 knockdown MCF10A cells transduced with control (MIG) or retrovirus to express the indicated EZH2 variants were analyzed by immunoblotting. GAPDH immunoblot served as a loading control. The expression levels of EZH2 in EZH2 knockdown (MIG) or EZH2 variant-expressing cells were controlled in **Supplementary Figure 5a**. Quantifications of Talin1 Rod domain (Talin1-R) were done by setting Talin1 full length (Talin1-F) plus Talin1-R signal as one and the ratios (Talin1-R/F+R) are shown in the upper bar graph. The relative abundance of head domain (Talin1-H) normalized to the level of Talin1-F is shown in lower bar graph. The amount of Talin1-H in EV control cells was set as 1. Data represent means ± s.e.m. from 3 independent experiments. Two-tailed Student's t-test with equal variance, * P = 0.018, ** P = 0.006, 0.009, 0.003, *** P = 0.0002, 3.3 x 10^{-7} (top to bottom). (**b**) BOSC23 cells were transfected with empty vector (EV) or plasmids to express the indicated EZH2 variants. Cytosolic extracts were isolated by lysing the cells directly on the culture plates to avoid any mechanical force-induced Talin cleavage and analyzed by immunoblotting. Talin1-R and Talin1-H levels were quantified as described in **a**. Data represent means ± s.e.m. from 3 independent experiments. Two-tailed Student's t-test with equal variance, * P = 0.01, ** P = 0.0002, 0.007, 0.003, *** P = 6.7 x 10^{-5}, 3.8 x 10^{-5} (top to bottom). (**c**) Talin1 from BOSC23 cells expressing EZH2 variants was immunoprecipitated and Talin1 methylation was determined by pan-methyl-lysine antibody. (n > 3). (**d**) Talin1 was immunoprecipitated from BOSC23, MDA-MD-231, or MCF10A cytosolic extracts and the association of EZH2 with VAV2 was determined by immunoblot using specific antibodies. * Background band detected by EZH2 antibody at approximately 120kDa. (n > 3). (**e**) Recombinant Talin1 (2300-2452) containing the C-terminal actin-binding site was purified and incubated with F-actin. Binding of Talin1 mutants to F-actin was determined using actin-co-sedimentation assays. Data

represent means ± s.d. (*n* =3). Two-factor ANOVA with replication, *** $P = 4.9 \times 10^{-31}$, 2×10^{-32} (left to right). (**f**) N-terminal GFP-fusion proteins of Talin1 variants (methyl-mimicking Talin1-K2454F or unmethylatable Talin1-K2454A or Q) transiently expressed in BOSC23 cells were detected by immunoblotting with anti-GFP antibody. Rod domain derived from cleaved GFP-Talin1 is indistinguishable from endogenous Rod domain and is not detectable by anti-GFP antibody. The bar chart shows the average percentage of cleaved head domain 'H' relative to full-length GFP-Talin1 'F' (*n* = 3). Data represent means ± s.d. Two-tailed Student's t-test, ** $P = 0.0008$, * $P = 0.03$. (**g**) Wild-type MCF10A cells were either left untreated or transduced with retrovirus to express wild-type EZH2 (EZH2), cytosolic EZH2 (EZH2ΔNLS), or VAV interaction mutant EZH2 (EZH2-VAV/MT). VAV2 knockdown was achieved by additional transduction with VAV2-specific shRNA-expressing lentivirus. Data shown are representative of 2 independent experiments. (**h**) H16N2 cells were transduced with retroviruses to express GFP-Talin1 variants. Adhesion structures and F-actin were visualized with anti-GFP antibody (green) and phalloidin (red), respectively. Representative images from one of more than 3 experiments are shown. Scale bar, 20 μm. Quantification of adhesion structures in H16N2 cells expressing GFP-Talin1 variants is shown in **Supplementary Figure 5f, g**. (**i**) H16N2 cells expressing Talin1 variants were analyzed as in **Figure 3c** and half-life values of adhesion assembly and disassembly were calculated based on GFP-Talin1 signal. Data represent means ± s.d. of 10-15 adhesions in 4-5 cells per group. Two-tailed student's t-test with equal variance, ** $P = 0.001$, *** $P = 5 \times 10^{-14}$, 2.5×10^{-8} (top to bottom) ($n > 3$).

Figure 5 PRC2 activity and sub-cellular distribution in EZH2 variant-expressing H16N2 cells

(**a**) H3K27me3 levels in H16N2 cells (as in **Figure 2a**) were determined by immunoblotting whole cell extracts. GAPDH served as a loading control. (n = 3) (**b**) EZH2 was immunoprecipitated from the cytosolic extracts of EZH2 knockdown MCF10A cells reconstituted with EZH2 variants and the association with SUZ12 and EED was determined by immunoblot analyses. *In vitro* methylation activity of immunoprecipitated EZH2 towards histone octamers was detected by H3K27me3-specific antibody. The sub-cellular fractionation was controlled in **Supplementary Figure 6**. (n > 3). (**c**) The interaction of EZH2 variants with CDYL and recruitment to EZH2 target genes was determined. EZH2 variants were immunoprecipitated from H16N2 cells to detect their interaction with CDYL (upper panel).

Chromatin immunoprecipitation (ChIP) from cells expressing EZH2 variants was performed using EZH2-specific antibody. Enrichments of EZH2 variants on the promoters of *ATF3* and *KDR*, compared to IgG control, were determined by qPCR. Data represents means ± s.d. of technical triplicates. Representative results from 2 independent experiments are shown. Asterisks in **b** and **c** indicate IgG controls using cytosolic extracts of EZH2ΔNLS-expressing cells. (**d**) Subcellular distributions of PRC2 components were analyzed by immunoblot analyses of whole cell, cytosolic, and nuclear extracts isolated from H16N2 cells expressing EZH2 variants. Proper sub-cellular fractionation was controlled by blotting for GAPDH (cytosol) and Lamin B (nucleus) (*n* = 3).

Figure 6 EZH2ΔNLS regulates the expression of STAT target genes

(**a**) Transcript levels of tumor suppressors reported to be regulated by EZH2 in various cancer cells were determined by RT-qPCR and normalized against GAPDH. Data represent means ± s.d. of 3 independent experiments with technical triplicates. (**b-c**) The differentially expressed genes between EZH2ΔNLS- and wild-type EZH2-overexpressing cells identified in RNA-seq experiments (full RNA-seq data sets are shown in **Supplementary Figure 7a, b** and source data for Figure 6) were analyzed to show the numbers (**b**) or enrichment *P*-values (**c**) of regulated genes that are targets of the indicated transcriptional regulators. Transcription factor binding sites within the potential transcriptional regulatory region (2 kb up-stream and 0.5 kb down-stream of the transcriptional start site) were determined using the ENCODE online tool (http://encodeproject.org/ENCODE/). Venn diagram in (**b**) shows the numbers of overlapping target genes of EZH2, MYC, STAT3 (direct), and the target genes of proteins up-regulated by STAT3 (STAT1, STAT2, c-Fos, and IRF1). Red bar in (**c**) indicates the enrichment *P* value of the differentially expressed EZH2 target genes among reported epithelial EZH2 targets. Target gene enrichment was determined using the Fisher's exact test (assuming all protein coding genes = 20,274). Detailed gene lists used for **b and c** are shown in the source file for Figure 6 and the raw data from the RNA-seq experiments that have been deposited to an online data base (ArrayExpress website). For reviewer's access see Methods.

Figure 7 EZH2-mediated Talin1 methylation promotes STAT3 activation and cellular transformation

(**a**) STAT3 immunoprecipitated from H16N2 cells expressing EZH2 variants (as in **Figure 2a**) was analyzed for methylation by pan-methyl-lysine antibody. Endo-EZH2-KD denotes knockdown of

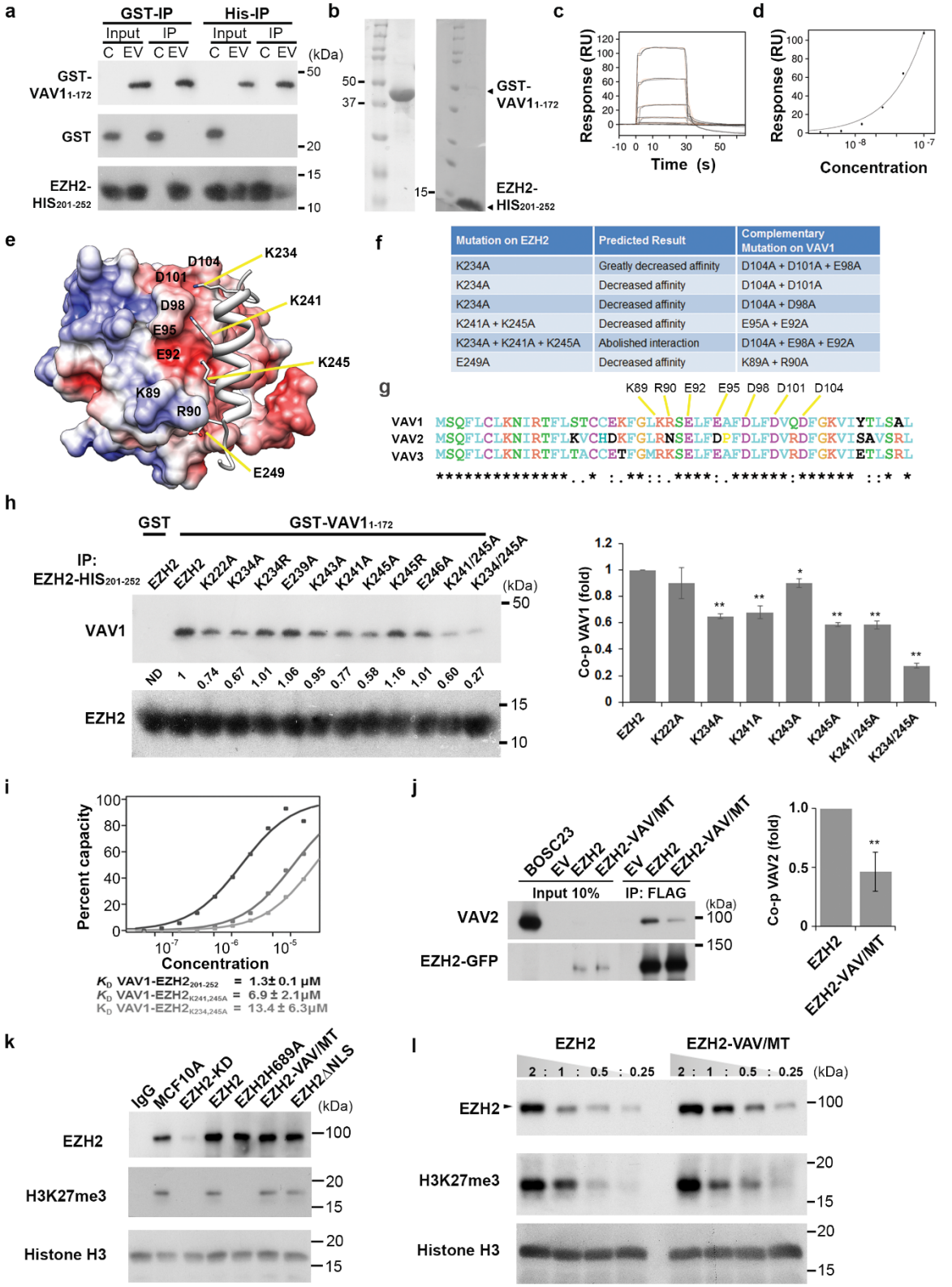
endogenous EZH2. Numbers below the Me-STAT3 immunoblot indicate the relative fold change after normalization to the immunoprecipitated STAT3 amounts. The abundance of Me-STAT3 in H16N2 cells was set as 1. Further quantifications are shown in **Supplementary Fig. 7c** (*n* = 3 or 4). (**b**) Immunoblots for phospho-STAT3 in extracts of H16N2 cells expressing EZH2 variants (treated with or without JAK2 inhibitor III SD-1029). The bar chart below the immunoblots shows the amounts of STAT3 pY705 in the indicated cells compared to the amounts observed in EZH2-expressing control cells after normalization to total STAT3. ND indicates not determined. Data represent means ± s.e.m. Two-tailed Student's t-test with equal variance, ** *P* = 0.0018, 0.0033, 0.0012 (top to bottom), *n* = 3. (**c**) Wild-type (MCF10A) or endogenous EZH2 knockdown MCF10A cells transduced with control virus (MIG) or retrovirus to express the indicated EZH2 variants were stimulated with or without IL-6 for 30 min. Whole cell extracts were subjected to immunoblotting with antibodies against STAT3 pY705 and total STAT3 (pY99). GAPDH served as a loading control. Data shown are representative of 2 independent experiments. (**d**) Immunoblots for phospho-STAT3 in extracts of H16N2 cells expressing Talin1 variants. The bar chart below the immunoblots shows the amounts of STAT3 pY705 in the indicated cells compared to the amounts observed in Talin1 expressing control cells after normalization to total STAT3. * *P* = 0.04, *n* = 4. (**e**) H16N2 cells overexpressing Talin1 variants were analyzed for anchorage independent cell growth as in **Figure 2a**, except that the colonies were also scored according to their size. Black: >60000 μm^2, light grey: 10000-60000 μm^2, dark grey: 4000-10000 μm^2. Data represent means ± s.d. (*n* = 3). Two-tailed Student's t-test with equal variance, * *P* = 0.03, *** *P* =0.0001, 0.00004, (top to bottom). (**f**) EZH2 knockdown H16N2 cells transduced with retrovirus to express the indicated EZH2 variants (EZH2, EZH2ΔNLS) or wild-type H16N2 cells un-transduced (H16N2) or transduced to express Talin1 variants (Talin1, Talin1-K2454F) were tested for anchorage independent cell growth. Cells were grown in medium untreated (un) or containing 200 μM of STAT3 inhibitor (S31-201) for 24 hours before culture in soft agar. Data represent means ± s.d. (*n* = 3, 3 independent experiments with technical triplicates). Two-tailed Student's t-test with equal variance, * *P* = 0.03, ** *P* = 0.0004, 0.0003, 0.0002, 0.01, *** *P* = 0.000097, 0.00004 (top to bottom). (**g**) Gene expression profiles of selected genes in H16N2 cells expressing EZH2 (blue shades) or Talin1 variants (green shades). The expression levels of exogenous EZH2 variants (Exo EZH2), Talin1 variants (GFP-Talin1), and indicated genes were measured by RT-qPCR or mean fluorescence intensity (MFI) of FACS analysis (Talin1 variants). Data represent means ± s.d. (*n* = 3). Two-tailed Student's t-test with equal variance, * *P* = 0.02, 0.02, 0.03, 0.04, ** *P* = 0.01,

0.0005, *** *P* = 0.00006 (top, left to right). * *P* = 0.02, ** *P* = 0.0007, 0.0005, 0.007, 0.003, 0.004, *** *P* = 0.00004 (bottom, left to right).

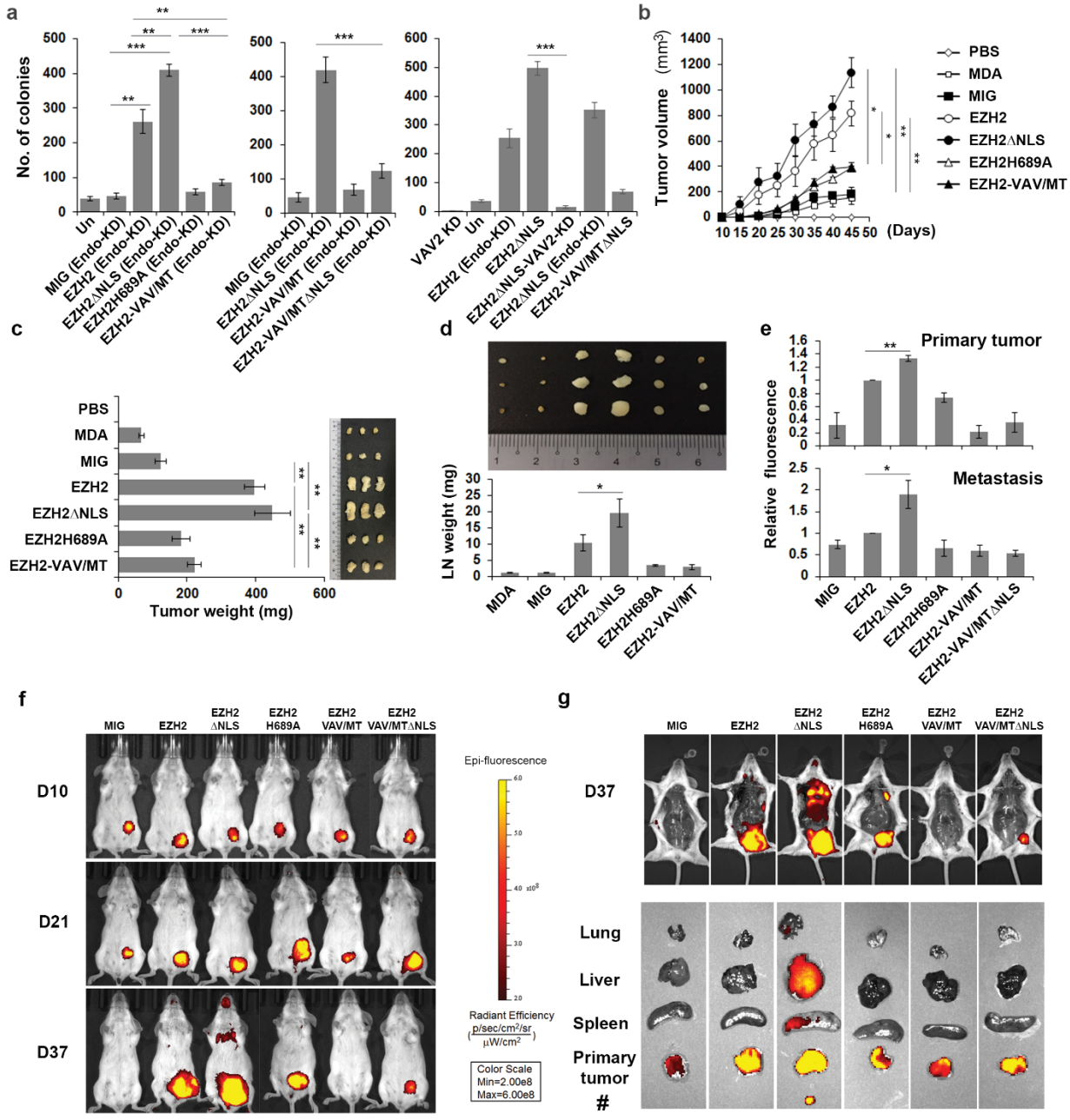
Figure 8 EZH2 is enriched in the cytoplasm of cancer stem cells

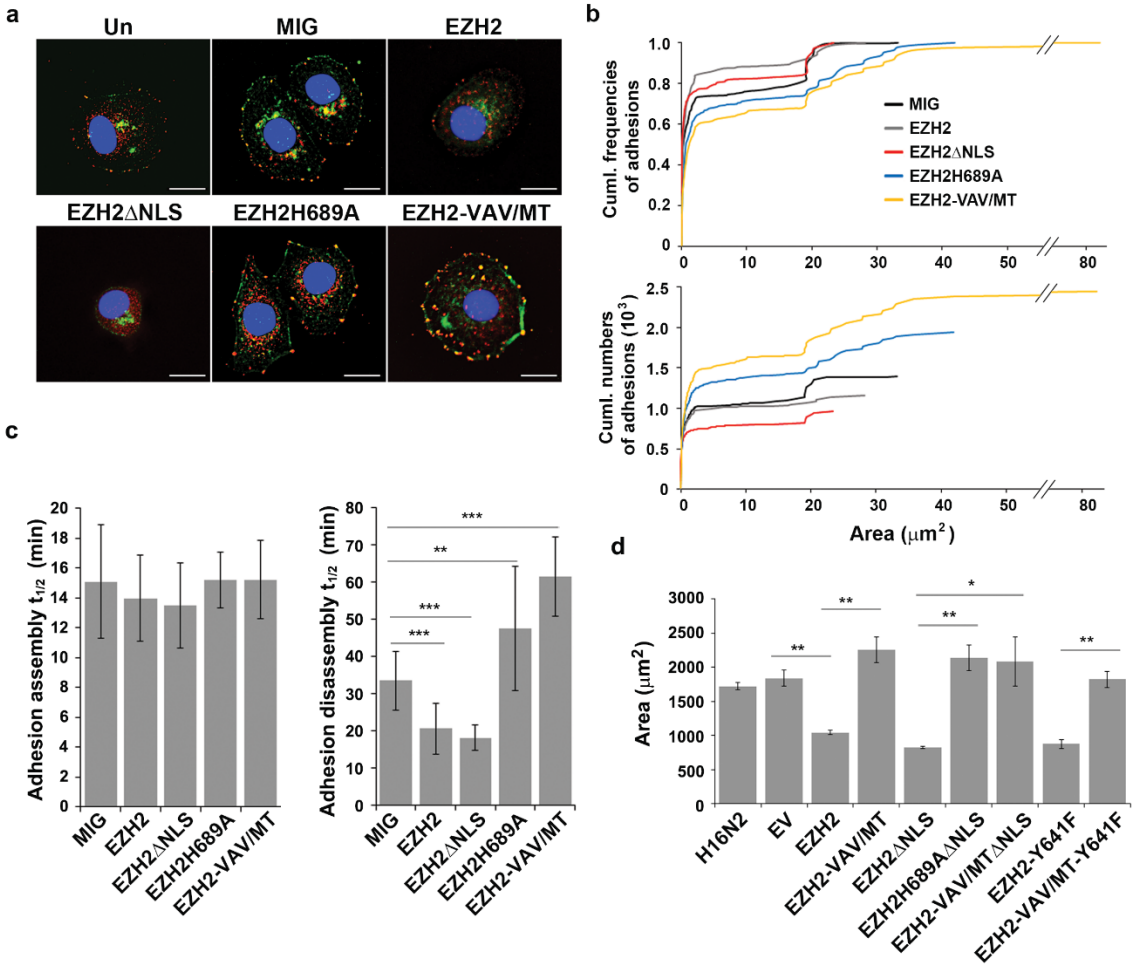
MDA-MB-231 cells were analyzed based on their surface expression levels of CD44 (high or low) and their ability to efflux Rhodamine 123 dye. CD44 high and Rhodamine 123 negative cells were defined as cancer stem cells with the side population (SP) phenotype ($CD44^{Hi}SP$). $CD44^{Hi}SP$ and CD44 low Rhodamine positive ($CD44^{Lo}Rd^{+}$) cells were also sorted from cultured MDA-MB-231 cells. EZH2 in the cells was visualized by immunofluorescent staining with specific antibody. (**a**) Representative 3D reconstructed z-stack confocal microscopy images from two independent experiments are shown. Scale bars, 10 μm. (**b**) Percentage of EZH2 in the cytoplasm was quantified based on the immunofluorescent signal of EZH2. Data represent means ± s.e.m. (*n* > 30 pooled from two independent experiments). *** *P* = 1.1 x 10^{-10}. (**c**) Subcellular distributions of EZH2 and STAT3 (phosphorylated and total) were analyzed by immunoblot analyses of whole cell, cytosolic, and nuclear extracts isolated from different subpopulations of MDA-MB-231 cells. Proper sub-cellular fractionation was controlled by blotting for GAPDH (cytosol) and Lamin B (nucleus).

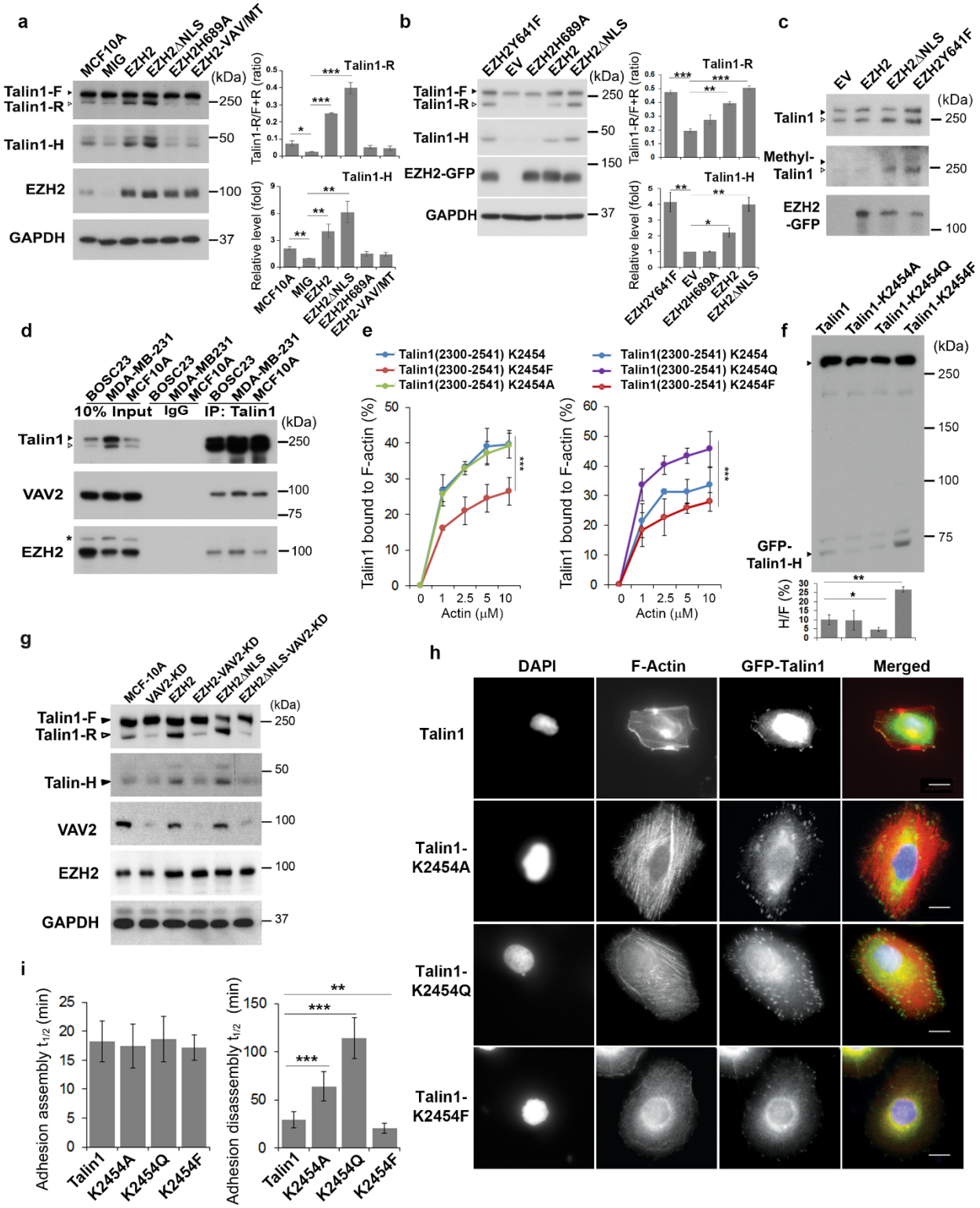
Venkatesan_Fig1

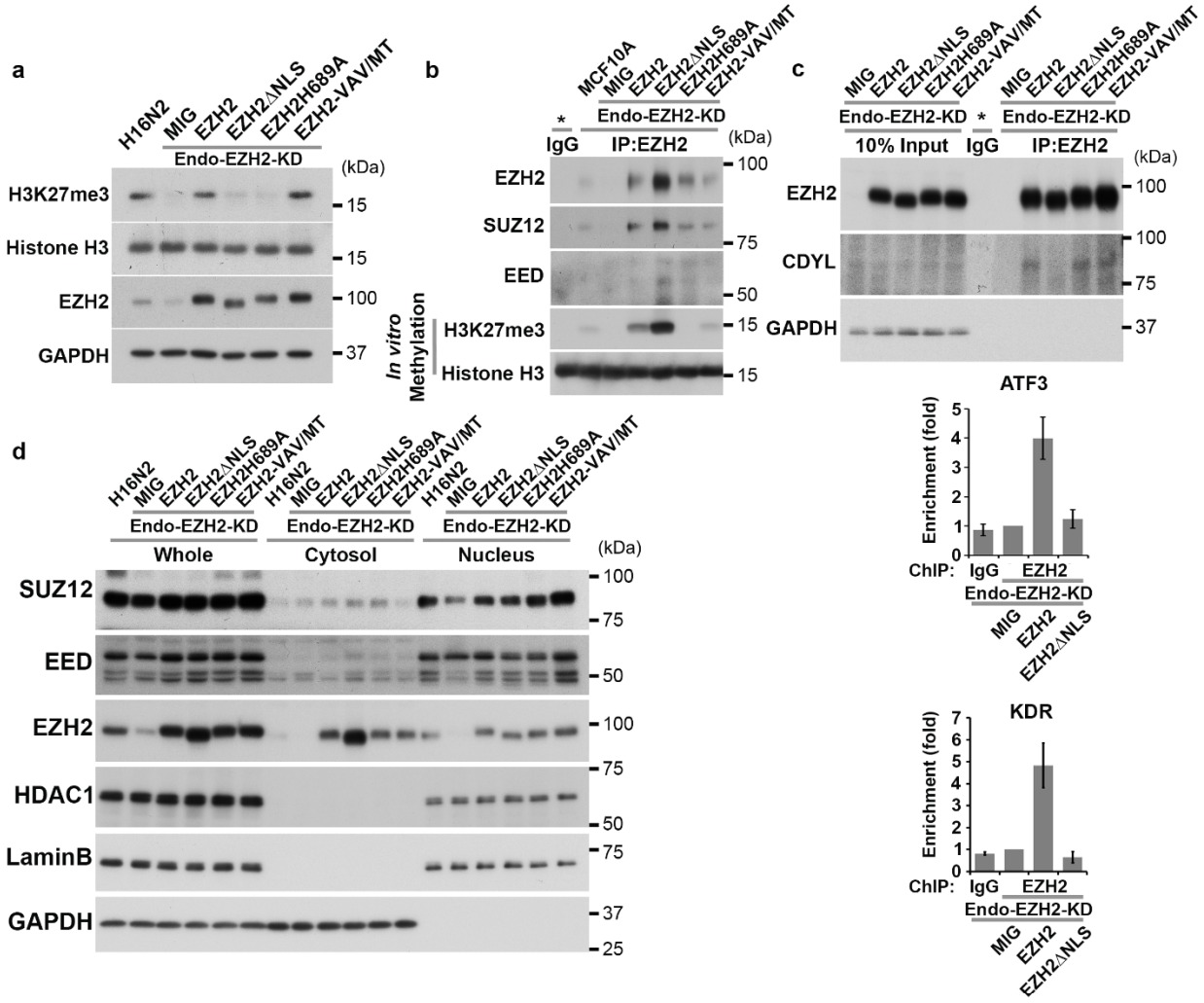


Venkatesan_Fig 2

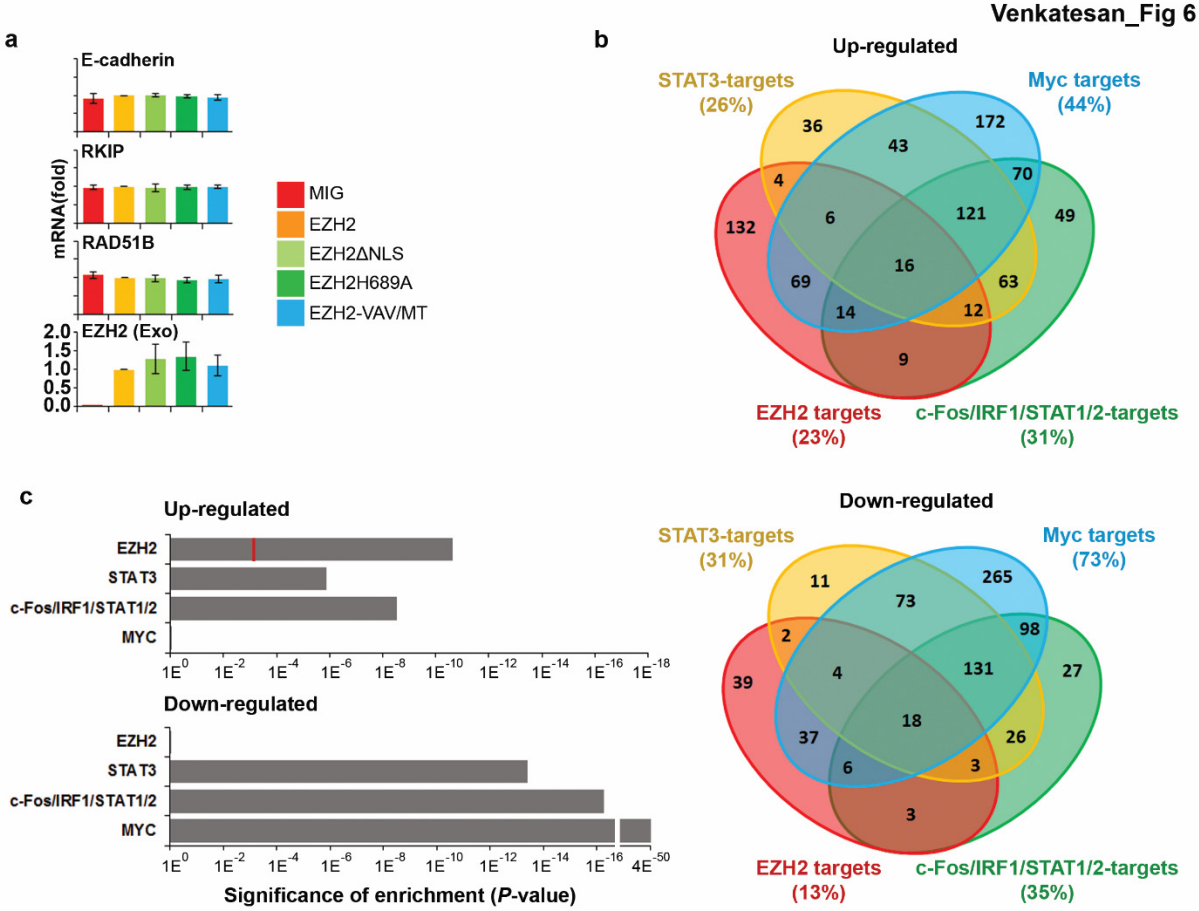


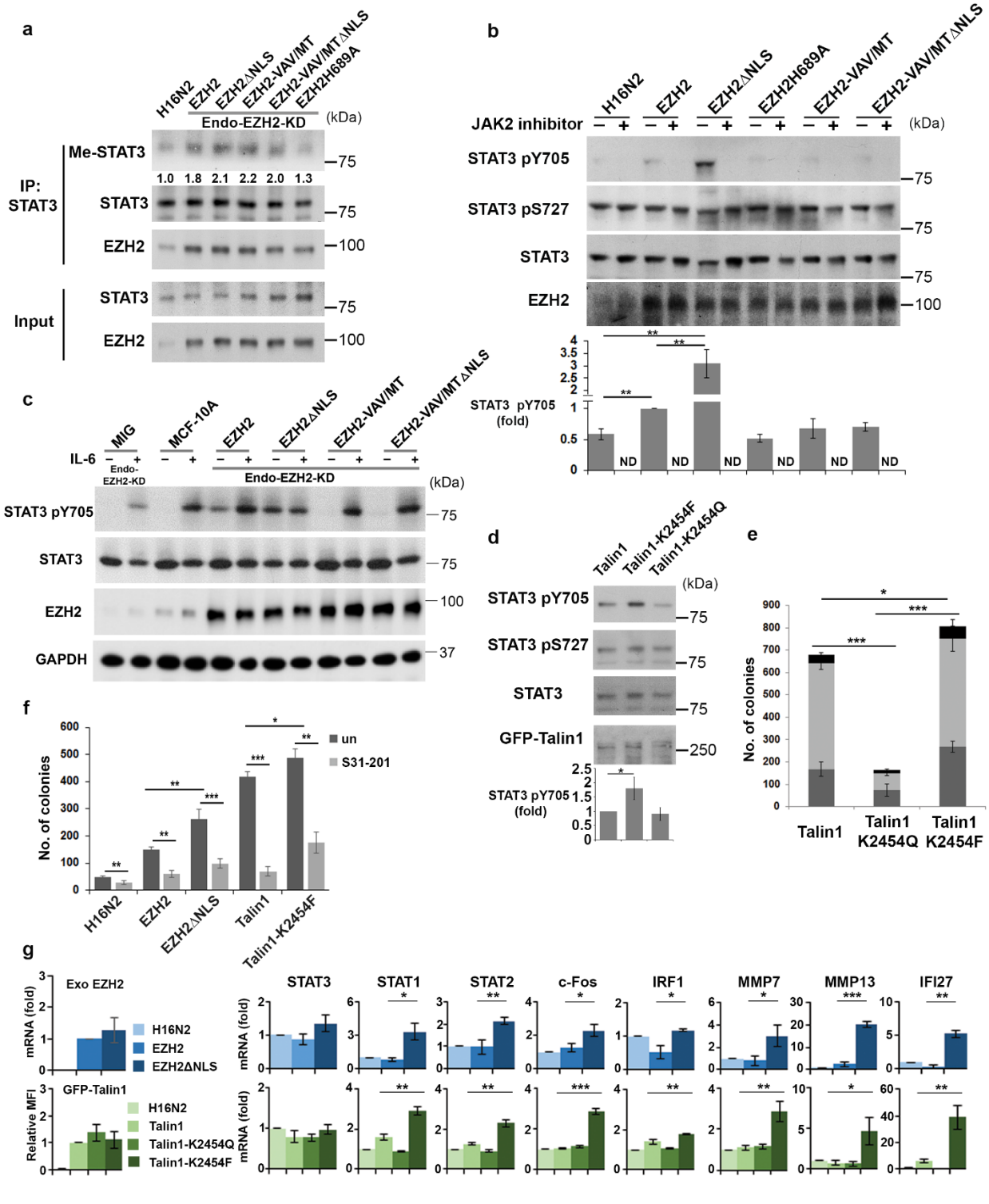

a
Un
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
b
Cuml. frequencies of adhesions
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
Cuml. numbers of adhesions (10^3)
Area (μm^2)
c
Adhesion assembly t1/2 (min)
Adhesion disassembly t1/2 (min)
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
d
Area (μm^2)
H16N2
EV
EZH2
EZH2-VAV/MT
EZH2ΔNLS
EZH2H689AΔNLS
EZH2-VAV/MTΔNLS
EZH2-Y641F
EZH2-VAV/MT-Y641F


a
MCF10A
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
(kDa)
Talin1-F
Talin1-R
Talin1-H
EZH2
GAPDH
250
50
100
37
Talin1-R
Talin1-R/F+R (ratio)
Talin1-H
Relative level (fold)
b
EZH2Y641F
EV
EZH2H689A
EZH2
EZH2ΔNLS
EZH2-GFP
150
c
Talin1
Methyl-Talin1
EZH2-GFP
d
BOSC23
MDA-MB-231
MCF10A
10% Input
IgG
IP: Talin1
VAV2
75
e
Talin1(2300-2541) K2454
Talin1(2300-2541) K2454F
Talin1(2300-2541) K2454A
Talin1(2300-2541) K2454Q
Talin1 bound to F-actin (%)
Actin (μM)
f
Talin1
Talin1-K2454A
Talin1-K2454Q
Talin1-K2454F
GFP-Talin1-H
H/F (%)
g
MCF-10A
VAV2-KD
EZH2
EZH2-VAV2-KD
EZH2ΔNLS
EZH2ΔNLS-VAV2-KD
Talin-H
h
DAPI
F-Actin
GFP-Talin1
Merged
Talin1
Talin1-K2454A
Talin1-K2454Q
Talin1-K2454F
i
Adhesion assembly t1/2 (min)
Adhesion disassembly t1/2 (min)

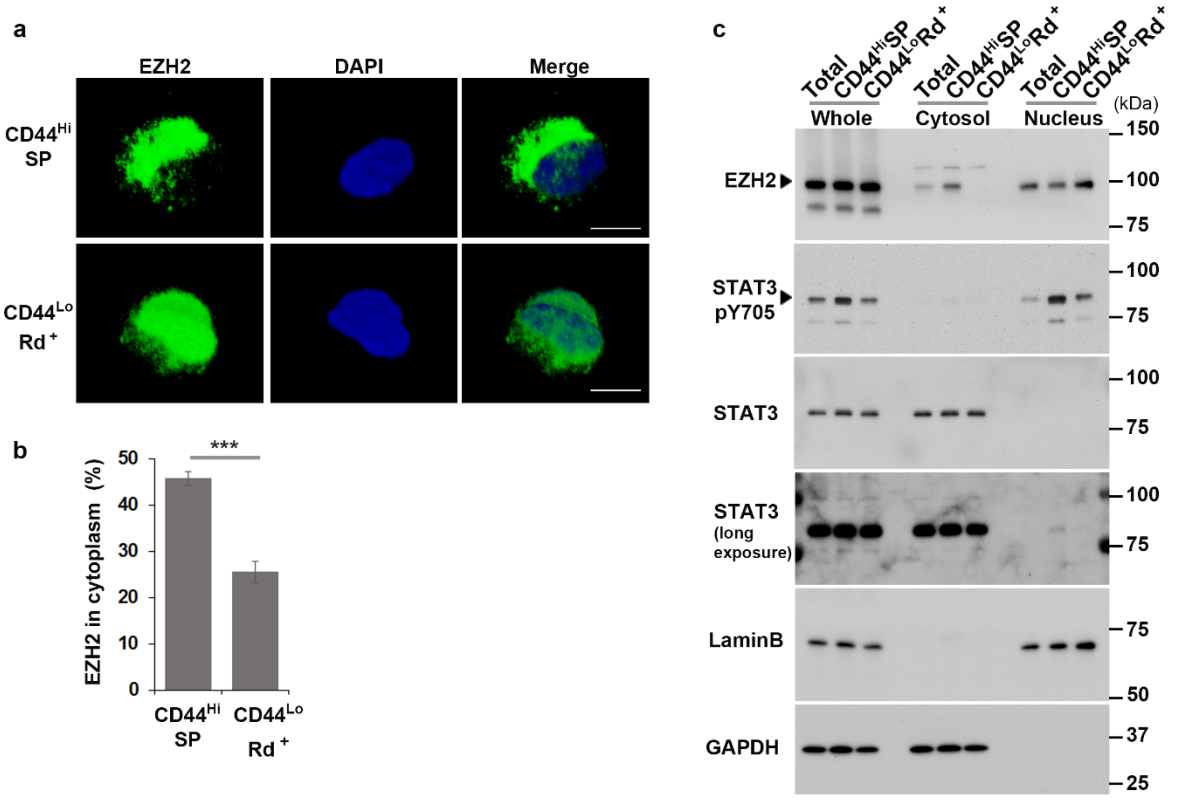

a
H16N2
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
Endo-EZH2-KD
(kDa)
H3K27me3
Histone H3
EZH2
GAPDH
15
15
100
37
b
MCF10A
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
Endo-EZH2-KD
IgG
IP:EZH2
(kDa)
EZH2
SUZ12
EED
H3K27me3
Histone H3
In vitro Methylation
100
75
50
15
15
c
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
Endo-EZH2-KD
10% Input
IgG
IP:EZH2
(kDa)
EZH2
CDYL
GAPDH
100
100
75
37
ATF3
Enrichment (fold)
ChIP:
IgG
EZH2
Endo-EZH2-KD
MIG
EZH2
EZH2ΔNLS
KDR
d
H16N2
MIG
EZH2
EZH2ΔNLS
EZH2H689A
EZH2-VAV/MT
Endo-EZH2-KD
Whole
Cytosol
Nucleus
(kDa)
SUZ12
EED
EZH2
HDAC1
LaminB
GAPDH
100
75
50
100
75
75
37
25

Venkatesan_Fig 6




a
H16N2
EZH2
EZH2ΔNLS
EZH2-VAV/MT
EZH2-VAV/MTΔNLS
EZH2H689A
Endo-EZH2-KD
(kDa)
IP: STAT3
Me-STAT3
1.0 1.8 2.1 2.2 2.0 1.3
STAT3
EZH2
Input
STAT3
EZH2
75
100
b
JAK2 inhibitor
STAT3 pY705
STAT3 pS727
STAT3
EZH2
STAT3 pY705 (fold)
ND
c
MIG
MCF-10A
IL-6
Endo-EZH2-KD
STAT3 pY705
STAT3
EZH2
GAPDH
37
d
Talin1
Talin1-K2454F
Talin1-K2454Q
GFP-Talin1
250
e
No. of colonies
Talin1
Talin1 K2454Q
Talin1 K2454F
f
un
S31-201
No. of colonies
H16N2
EZH2
EZH2ΔNLS
Talin1
Talin1-K2454F
g
Exo EZH2
mRNA (fold)
GFP-Talin1
Relative MFI
STAT3
STAT1
STAT2
c-Fos
IRF1
MMP7
MMP13
IFI27

Venkatesan_Fig 8



EZH2 promotes neoplastic transformation through VAV interaction-dependent extranuclear mechanisms

Nandini Venkatesan[1,7], Jong Fu Wong[1,7], Kuan Pern Tan[2,6], Hwa Hwa Chung[1], Yin Hoe Yau[1], Engin Cukuroglu[5], Abdollah Allahverdi[1], Lars Nordenskiold[1], Jonathan Göke[5], Susana Geifman Shochat[1], Valerie Chun Ling Lin[1], M.S. Madhusudhan[2,3,4], I-hsin Su[1*]

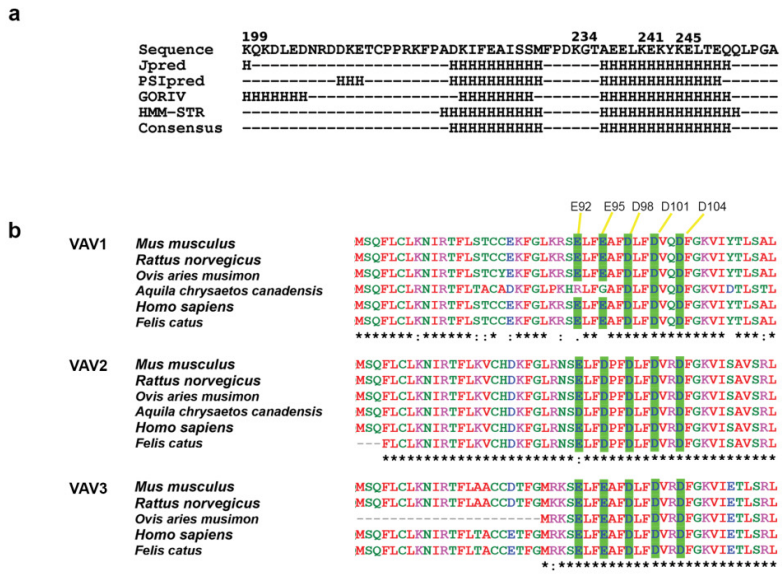
Supplementary information

1. Supplementary Figures and Figure Legends

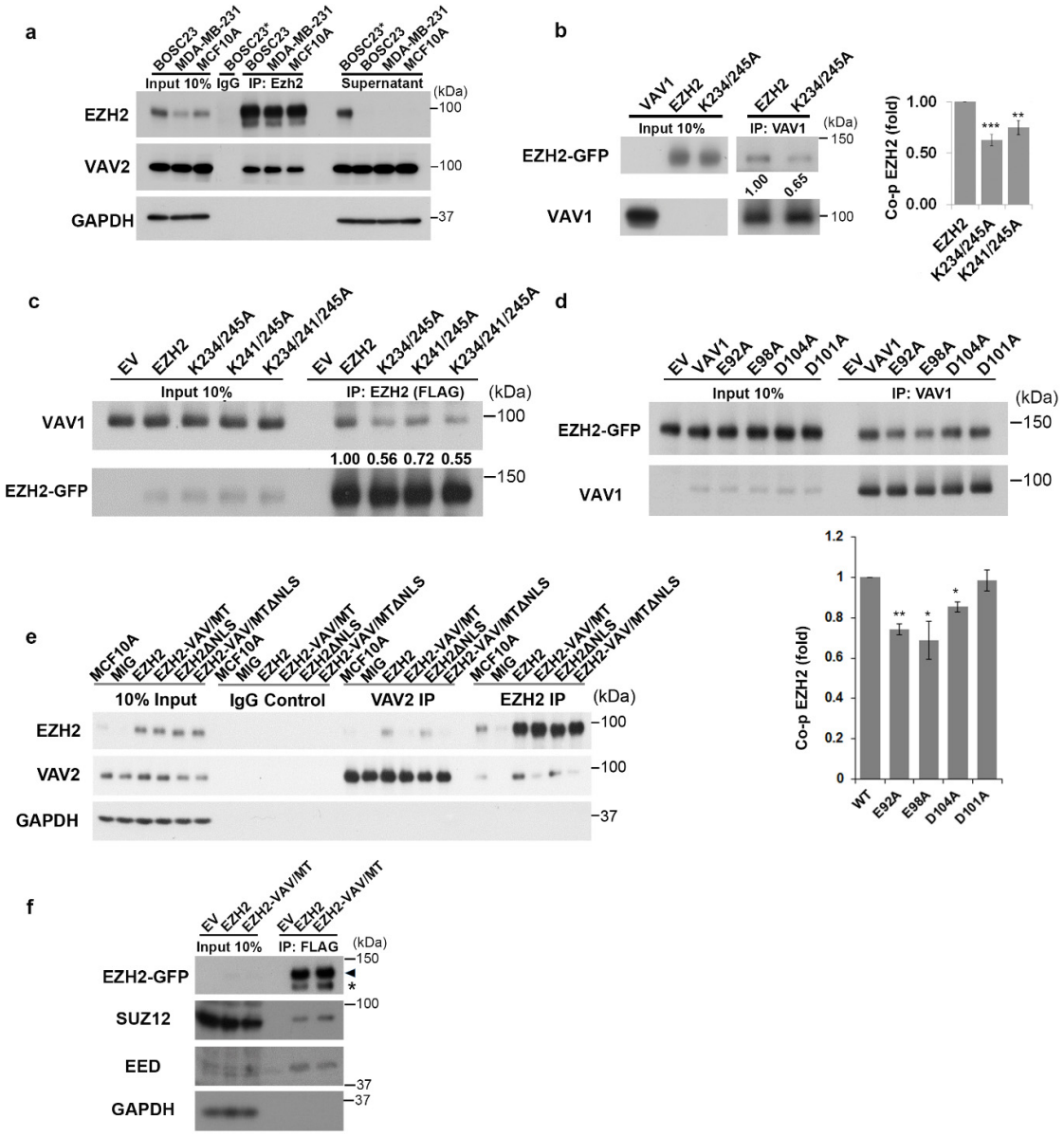
2. Supplementary Videos

3. Supplementary Methods

1. Supplementary Figures and Figure Legends



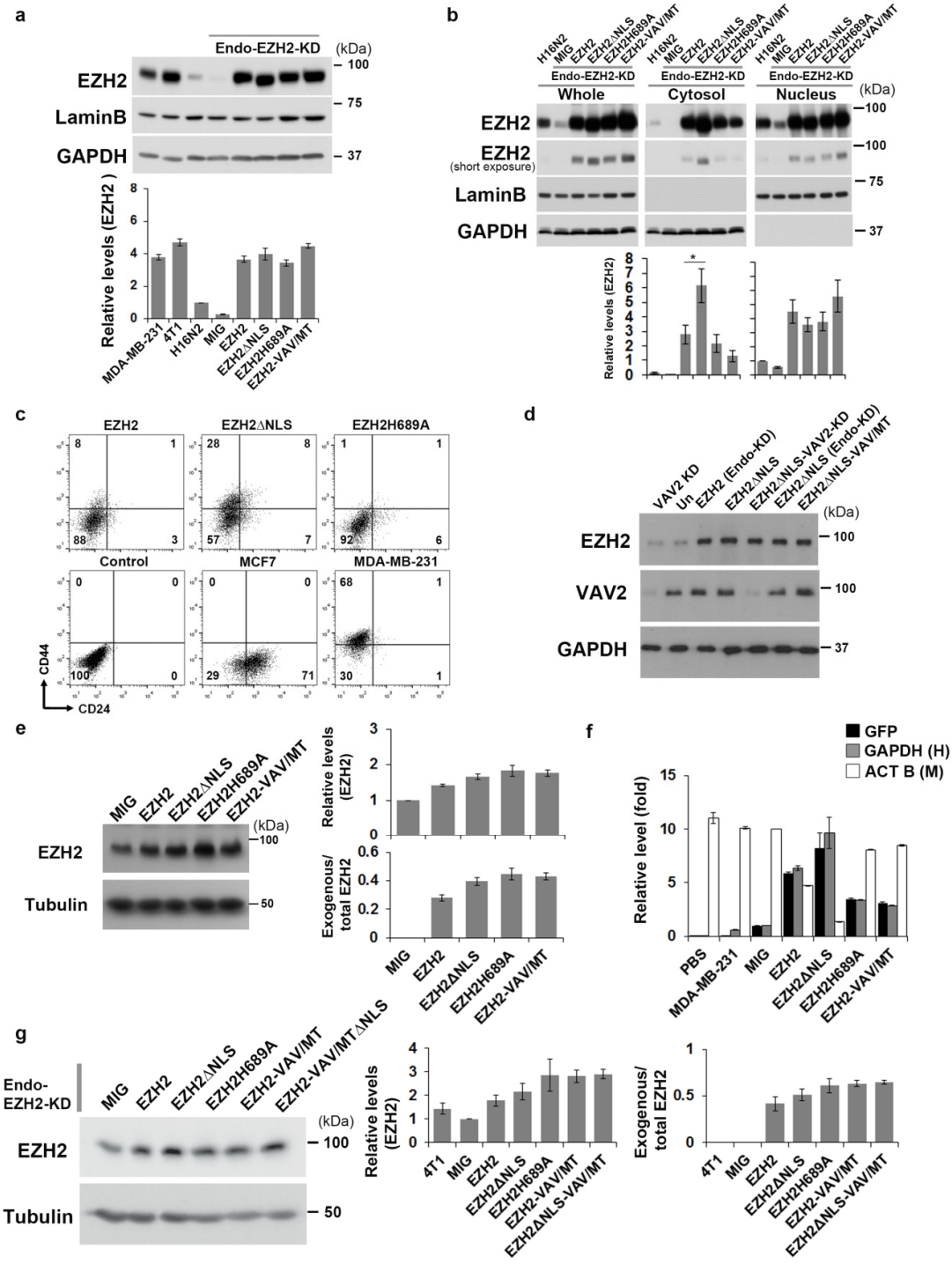
Supplementary Figure 1 Conservation of the EZH2 interacting domain in VAV family proteins. (**a**) Secondary structure prediction for the EZH2 protein sequence that interacts with VAV1 using Jpred, PSIpred, GORIV, and HMM-STR. The consensus (by majority) of these methods (last line) predicts 2 helices between residues 221-230 and 237-250. (**b**) Sequence alignment of VAV family proteins from the indicated species was performed using the ClustalW2 online tool (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The conserved acidic amino acids are highlighted in green. Asterisks indicate positions that have a single fully conserved residue, colons indicate conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet PAM 250 matrix). Period indicates conservation between groups of weakly similar properties - scoring ≤ 0.5 in the Gonnet PAM 250 matrix.



Supplementary Figure 2 EZH2 interacts with VAV2 and the VAV interaction mutant EZH2 associates with core components of PRC2.

(**a**) Cytosolic extracts were isolated from the indicated human cell lines and EZH2 was immunoprecipitated using specific antibody. Co-immunoprecipitation of VAV2 with EZH2 was determined by immunoblotting with specific antibodies. GAPDH is the loading control. Data shown are representative of 3 independent experiments. * Cytosolic extract of BOSC23 cells was used for IgG immunoprecipitation as a negative control. (**b**) Endogenous VAV1 immunoprecipitated from Jurkat whole cell extracts was incubated with equal amounts of wild-type EZH2 or EZH2-K234/245A mutant protein isolated from transfected BOSC23 cells to test the interaction between full length VAV1 and EZH2. The amount of wild-type EZH2 coprecipitated with VAV1 was set to 1. Data represent means ± standard deviation (s.d.) of 4 independent experiments. Two-sided Student's t-test with equal variance, $**P = 0.0004$, $***P$

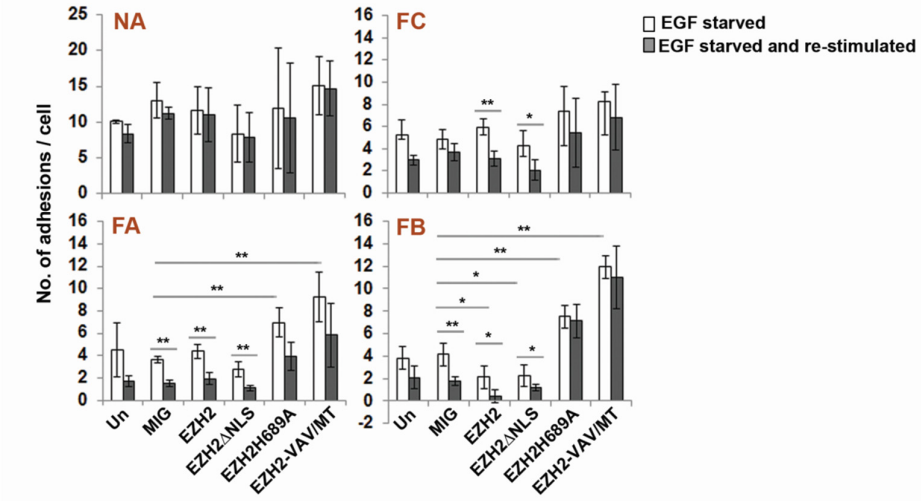
= 1.2 x 10^{-5}. (**c**) Flag-tagged EZH2-GFP mutants were purified by FLAG immunoprecipitation from BOSC23 cells transfected with EZH2 mutant constructs or empty vector (EV). VAV1 was purified from Jurkat lysate by VAV1 IP and eluted using VAV1 peptide. VAV1 was incubated overnight with EZH2-GFP on beads or control beads (EV). Numbers below VAV1 blot indicate the relative amount of VAV1 interacting with mutant EZH2 compared to that with wild-type EZH2. (**d**) Flag-tagged EZH2-GFP was purified by FLAG immunoprecipitation from transfected BOSC cells. VAV1 mutants were purified from BOSC23 cells transfected with mutant VAV1 constructs or empty vector (EV) by VAV1 IP. Purified EZH2-GFP was incubated overnight with VAV1 mutants on beads or control beads (EV). Bar graph below summarizes the fold changes of co-precipitated EZH2 (Co-p EZH2) by VAV mutants relative to that pulled down by wild-type VAV1 from 3 independent experiments. Data represent means ± s.d. One-sided Student's t-test, ** P = 0.002, * P =0.01, 0.02 (left to right). (e) Whole cell extracts of wild-type (MCF10A) or endogenous EZH2 knockdown MCF10A cells transduced with control virus (MIG) or retrovirus to express the indicated EZH2 variants were subjected to EZH2 or VAV2 immunoprecipitation. Data shown are representative of 2 independent experiments. (**f**) Whole cell extracts obtained from BOSC23 cells transiently expressing EZH2 flag-tagged variants were subjected to anti-Flag immunoprecipitation, coprecipitated endogenous SUZ12 and EED were determined by immunoblotting with specific antibodies. Triangle and asterisk indicate the position of EZH2-GFP and a degradation product of the fusion protein, respectively. GAPDH served as a loading control. Data shown are representative of 3 independent experiments.



Supplementary Figure 3 Expression of exogenous EZH2 variants in mammary epithelial and breast cancer cells.

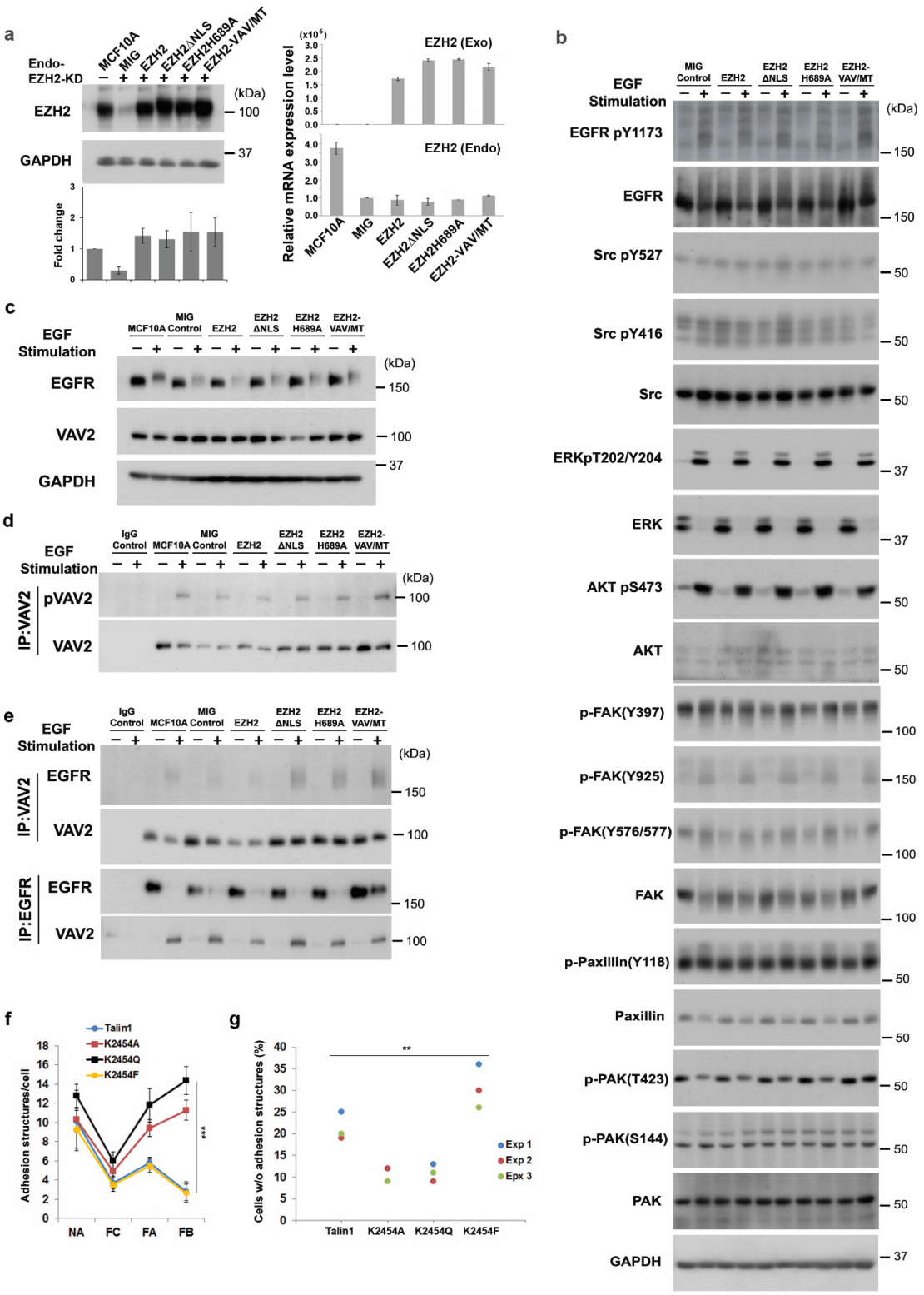
(**a**) Expression levels of EZH2 variants. Whole cell lysates of MDA-MB-231 (MDA), 4T1, H16N2, and endogenous EZH2 knockdown H16N2 (Endo-EZH2-KD) cells overexpressing EZH2 variants were analyzed by immunoblot. LaminB and GAPDH were used as

normalization controls to determine the relative EZH2 levels among the different cell lines. Bar graph summarizes the quantification of normalized results from 3 independent experiments and data represent means ± standard error of the mean (s.e.m.). The EZH2 amount in H16N2 cells was set as 1. (**b**) Subcellular localization of EZH2 variants. Cytoplasmic and nuclear proteins extracted from mammary epithelial cells (H16N2 or H16N2-EZH2-KD cells expressing the indicated EZH2 variants) were subjected to immunoblot analysis with specific antibodies. Proper sub-cellular fractionation was controlled by blotting for GAPDH (cytosol) and Lamin B (nucleus). Data shown are representative of 4 independent experiments and the normalized quantification is shown in the bar graphs below. Data shown are means ± s.e.m. Two-tailed Student's t test, * $P = 0.04$. The endogenous nuclear EZH2 amount in H16N2 cells was set as 1. (**c**) Cancer stem cell markers on cells expressing EZH2 variants. H16N2 cells, in which endogenous EZH2 was knocked down, were transduced with retroviruses to express the indicated EZH2 variants and examined for the surface expression of CD44 and CD24. Stained parental cells were used as a negative control. MCF7 and MDA-MB-231 were positive controls for CD24 and CD44, respectively ($n = 3$). (**d**) Wild-type un-transduced (Un), endogenous EZH2 knockdown (Endo-KD) H16N2 or VAV2 knockdown (VAV2-KD) cells transduced with control (MIG) or retrovirus to express the indicated EZH2 variants were analysed by immunoblot and used for anchorage independent cell growth in Figure 2a, right panel. (**e**) EZH2 expression levels and ratios of exogenous EZH2 in MDA-MB-231 cells transduced with control virus (MIG) or retroviruses for expression of the indicated EZH2 variants. Tubulin is the loading control. Bar graph summarizes results from 3 independent experiments and normalized data are represented as means ± s.e.m.. The EZH2 amount in control virus (MIG) transduced MDA-MB-231 cells was set as 1 (upper panel). The ratio of exogenous EZH2 was calculated by subtracting the level of EZH2 in MIG and dividing by the total EZH2 level in cells (bottom panel). (**f**) cDNA was generated using RNA extracted from the popliteal lymph nodes of mice received xenografted MDA-MB-231 cells, as shown in Fig. 2d. Transcript levels of GFP (co-expressed with exogenous EZH2 in MDA-MB-231 cells), human GAPDH (GAPDH (H)) and mouse beta-actin (ACT B (M)) were determined using real-time PCR. A region of 18S ribosomal RNA conserved between human and mouse was used as the loading reference. Transcript levels in MIG were set as 1 (for GFP, GAPDH) or 10 (for ACT B). Data shown is means ± s.d. of technical triplicates. (**g**) EZH2 expression levels and the ratio of exogenous EZH2 in 4T1 cells. Endogenous EZH2 was reduced in 4T1 cells by shRNA-mediated knockdown, as revealed in MIG control cells. Expression levels of the indicated EZH2 variants in EZH2 knockdown 4T1 cells were visualized by immunoblotting for EZH2. Tubulin was used as a loading control. Bar graphs summarize the normalized results from 3 independent experiments. Data are represented as means ± s.e.m. EZH2 amount in EZH2 knockdown cells (MIG) was set as 1 (left panel). The ratio of exogenous EZH2 was calculated by subtracting the level of EZH2 in MIG and dividing by the total EZH2 level in cells (right panel).



Supplementary Figure 4 EGF stimulated changes in adhesion structures are not regulated by EZH2.

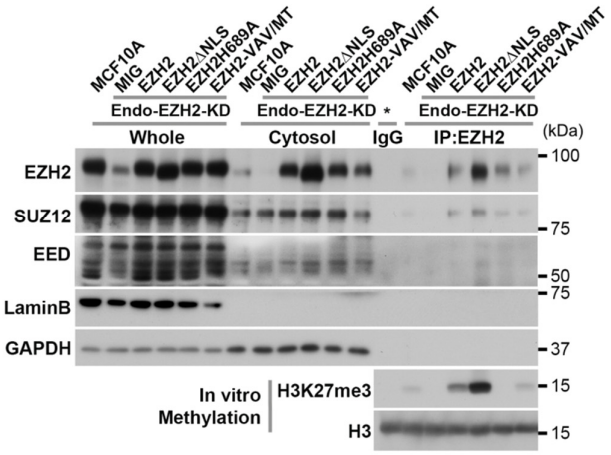
H16N2 cells were transduced with control virus (MIG) or retroviruses for the expression of the indicated EZH2 variants and subsequently endogenous EZH2 was knocked down by shRNA. "Un" indicates un-transduced wild-type control cells. Cells were EGF starved for 3 days and re-stimulated with EGF for 12 hours. Adhesion structures were visualized using a specific antibody against Paxillin. Bar graphs show a summary of three independent experiments. Nascent adhesions (NAs), focal complexes (FCs), focal adhesions (FAs) and fibrillar adhesions (FBs) were measured by setting the size limits of the measured particles from 0.05-200 μm^2. The classifications are as follows: NA $\leq$ 0.2 μm^2, 0.2 μm^2 < FC $\leq$ 0.8 μm^2, 0.8 μm^2 < FA $\leq$19 μm^2, FB > 19 μm^2. Data represent means ± s.d. of 3 independent experiments with 50-60 cells per group per experiment scored. One-tailed Student's t-test with equal variance, FC: ** P = 0.004, * P = 0.041, FA: ** P = 0.006, 0.007, 0.0036, 0.00035, 0.007 (top to bottom), FB: * P = 0.019, 0.019, 0.013, 0.023 (top to bottom). ** P = 0.007, 0.01, 0.007 (top to bottom).



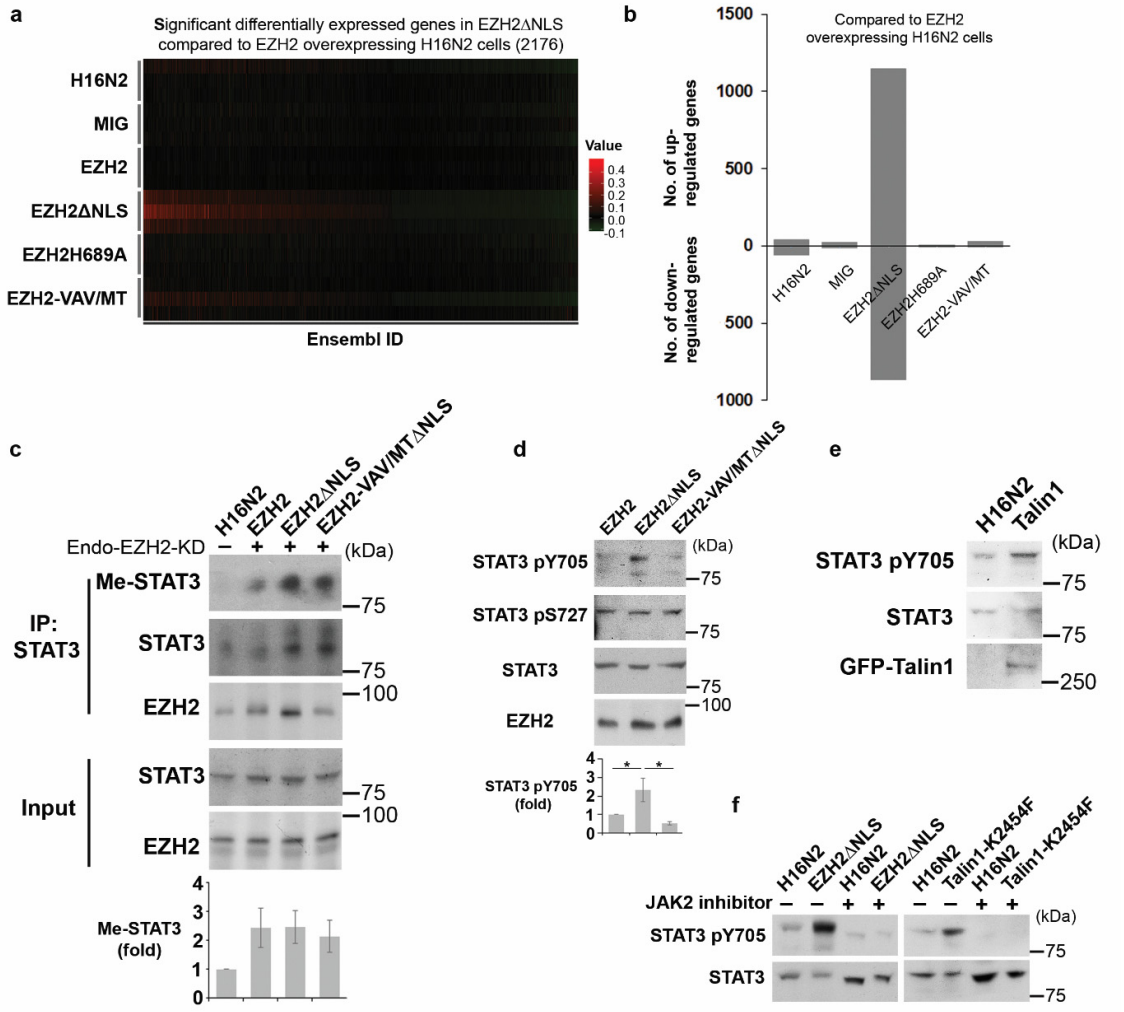
Supplementary Figure 5 EZH2 does not control EGF signaling, but regulates the formation of prominent adhesion structures through Talin1 methylation.

(**a**) Endogenous EZH2 in human mammary epithelial MCF10A cells was eliminated by shRNA-mediated knockdown. Protein levels were determined by immunoblot (left upper

panel). The left lower bar graph summarizes the normalized immunoblot results from 3 independent experiments and data represent means ± s.e.m. The mRNA expression levels of endogenous EZH2 (right lower bar graph) and exogenous EZH2 variants (right upper bar graph) were determined by qRT-PCR. (**b**) MCF10A cells expressing EZH2 variants as in (**a**) were EGF starved for 24 hours and re-stimulated for 30 minutes before harvesting whole cell lysates. Activation of the indicated proteins was determined by immunoblot analyses using specific antibodies. GAPDH was used as a loading control. (**c-e**) VAV2 activation and recruitment to EGFR. MCF10A cells expressing EZH2 variants as in (**a**) were starved and stimulated as described in (**b**). The total protein levels of EGFR and VAV2 are shown in (**c**). (**d**) EGF-stimulated VAV2 phosphorylation was determined by immunoblotting with anti-phospho-VAV antibody after VAV2 immunoprecipitation. (**e**) EGF-stimulated VAV2-EGFR interaction was determined by reciprocal immunoprecipitation. Data shown in (**a-e**) are representative of more than 3 independent experiments. (**f**) H16N2 cells were transduced with retroviruses to express wild-type GFP-Talin1, GFP-Talin1-K2454A, Q or F. Various adhesion structures were visualized by fluorescence microscopy after immunostaining with anti-GFP antibody and images were analyzed using the "Measure/Analyze Particle Tool" in ImageJ. Quantification of H16N2 cells expressing Talin1 variants with various adhesion structures was performed as defined in S**upplementary Figure 4.** Data represent means ± s.d. of 3 independent experiments with around 100 cells per group. (**g**) Quantification of H16N2 cells expressing Talin variants without visible adhesion structures. All data points are shown from 3 independent experiments with around 100 cells per group. Statistics among groups were performed using two-way ANOVA, *** P = 2.3E-10 (**f**), Single-factor ANOVA ** P = 0.002 (**g**).



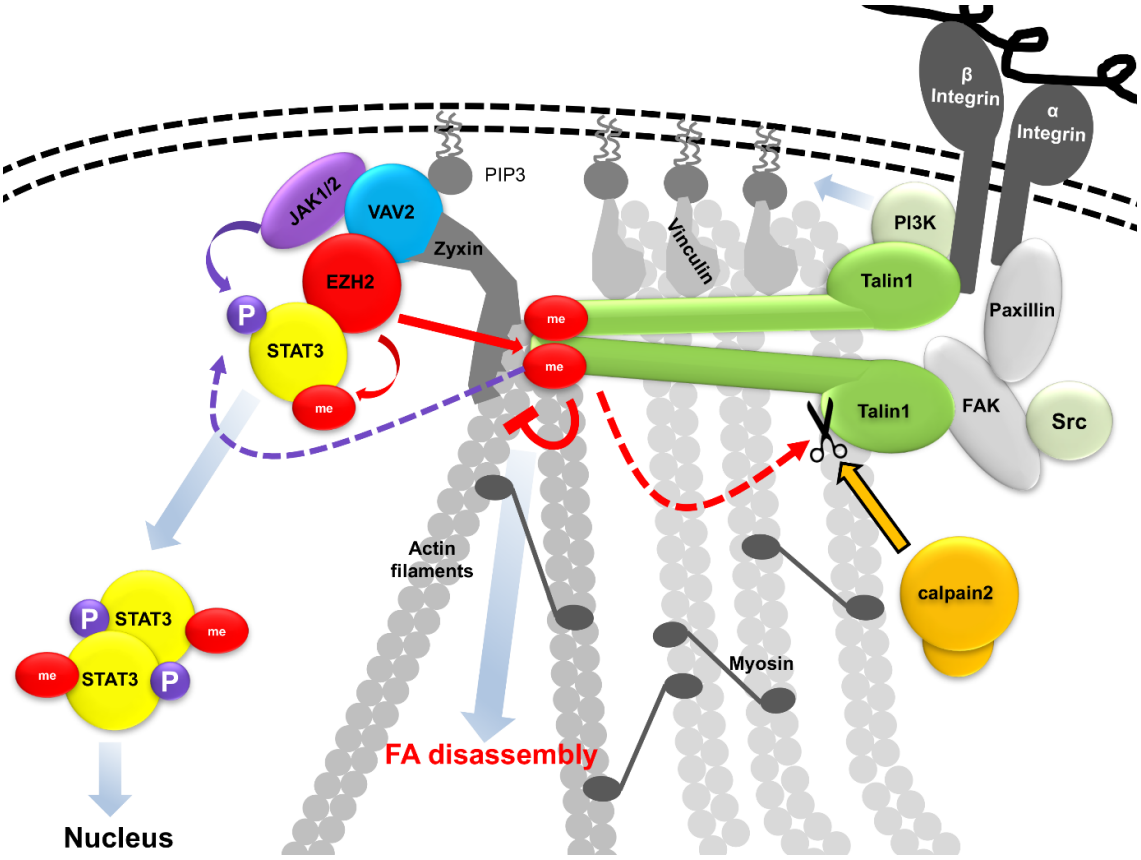
Supplementary Figure 6 Deletion of the nuclear localization signal in EZH2 did not compromise PRC2 complex formation and *in vitro* methylation activity of EZH2.
EZH2 immunoprecipitated from cytosolic extracts of MCF10A cells (as in **Figure 5b**) was analysed in immunoblots to determine the co-immunoprecipitation of SUZ12 and EED. Whole cell and cytosolic extracts were included to control the quality of fractionation (Controls for **Figure 5b**). * Cytosolic extract of EZH2ΔNLS expressing MCF10A cells was used for IgG immunoprecipitation as a negative control.



Supplementary Figure 7 EZH2ΔNLS and methyl-Talin1 regulate the expression of STAT3 target genes.

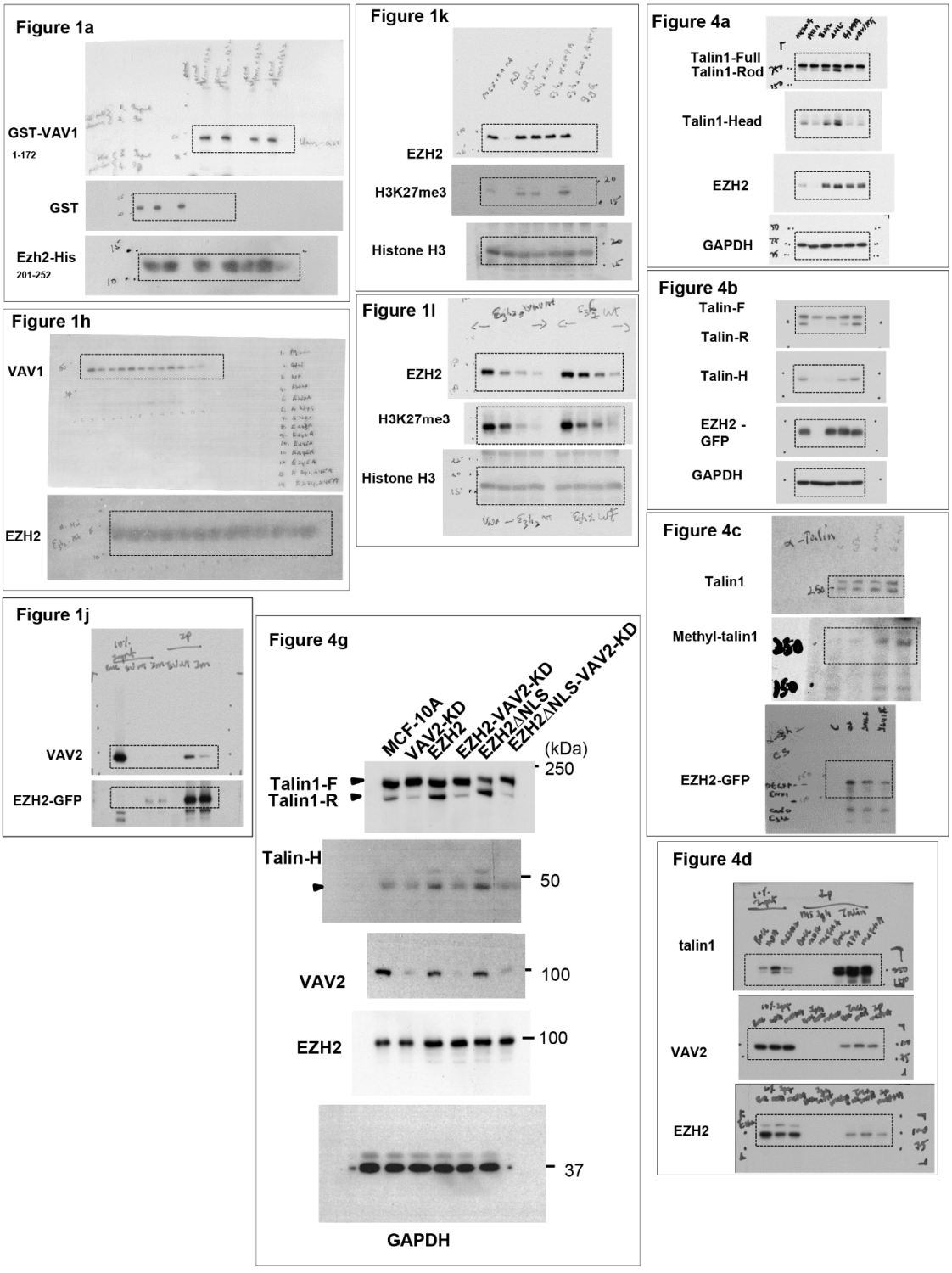
(**a**,**b**) RNA-seq results from H16N2 cells (as in **Figure 2a**) are summarized in a heat-map. Normalized FPKM values for differentially expressed genes in each group were compared to those in EZH2 expressing cells. The data were normalized such that the sum of the expression estimates for each gene was 1 (relative expression) and the normalized FPKM value from the EZH2 sample was subtracted from the other samples to show the EZH2 sample as the baseline. Three independent sets of RNA samples derived from 2 independent transductions were included in this analysis. (**b**) Total numbers of differentially expressed genes identified in the RNA-seq analysis of EZH2 variant-expressing H16N2 cells are summarized. Detailed gene lists used for this figure are shown in the source file for Figure 6 and the raw data of RNA-seq experiments that have been deposited to online an data base (ArrayExpress website). For reviewer's access please see Methods. (**c**) STAT3 immunoprecipitated from H16N2 cells expressing EZH2 variants was analyzed for methylation by pan-methyl-lysine antibody. ENDO-EZH2-KD denotes knockdown of endogenous EZH2. (**d**,**e**) Immunoblots for phospho-STAT3 in extracts of H16N2 cells expressing EZH2 variants (**d**) or Talin1 (**e**). The bar charts show the average amounts of methylated STAT3 (**c**) or phospho-STAT3 (pY705) (**d**) in the indicated cells compared to the amounts observed in control cells (**c**) or EZH2 expressing cells (**d**) (n = 3). Data represent means ± s.d. Two-tailed Student's t test, * P = 0.05, 0.02 (left to

right). (**f**) Phospho-STAT3 levels in control and H16N2 cells expressing EZH2ΔNLS or Talin1-K2454F (with or without JAK2 inhibitor III SD-1029) were determined by immunoblot ($n = 3$).



Supplementary Figure 8 Graphic Summary
EZH2 in the cytoplasm methylates Talin, thereby promoting adhesion turnover through disruption of Talin binding to F-actin and enhancement of calpain 2-dependent Talin cleavage. Meanwhile, STAT3 is methylated and recruited to VAV by EZH2, leading to VAV-interaction-dependent JAK2-mediated phosphorylation and activation of STAT3. The dashed lines with arrowheads indicate processes promoted by methylated Talin1 through unknown mechanisms. We propose that these events contribute to EZH2-promoted cellular transformation in mammary epithelial cells.

Supplementary Figure 9 Uncropped immunoblots



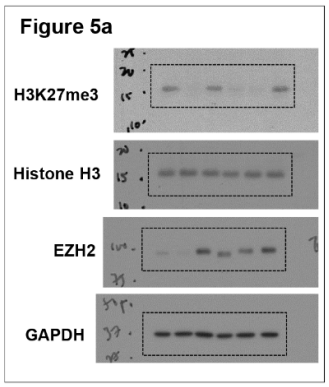


Figure 5a
H3K27me3
Histone H3
EZH2
GAPDH

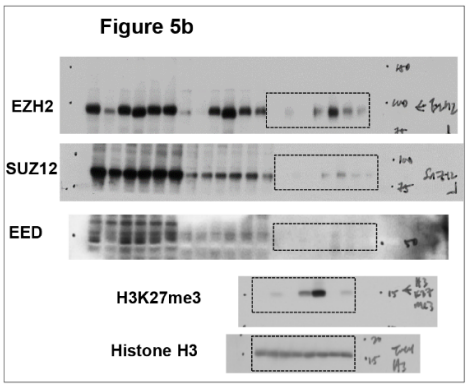


Figure 5b
EZH2
SUZ12
EED
H3K27me3
Histone H3

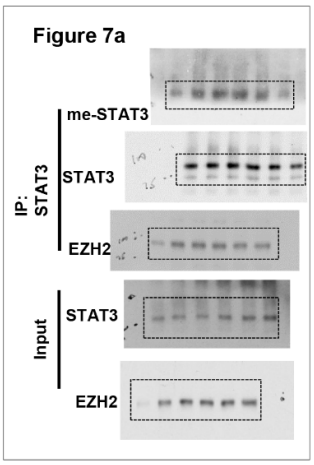


Figure 7a
IP: STAT3
me-STAT3
STAT3
EZH2
Input
STAT3
EZH2

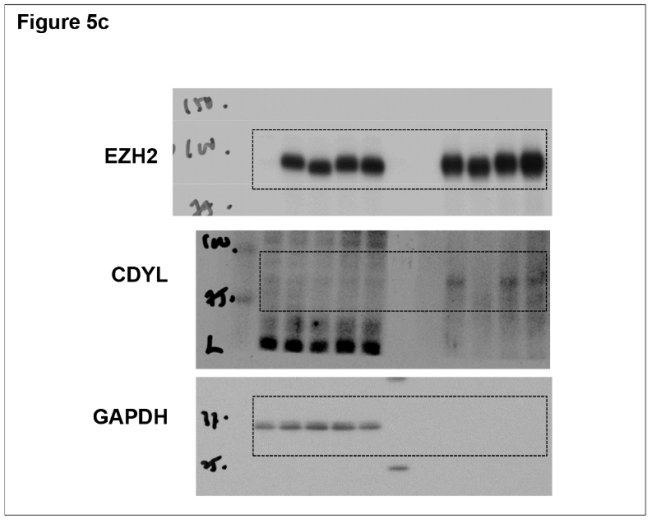


Figure 5c
EZH2
CDYL
GAPDH

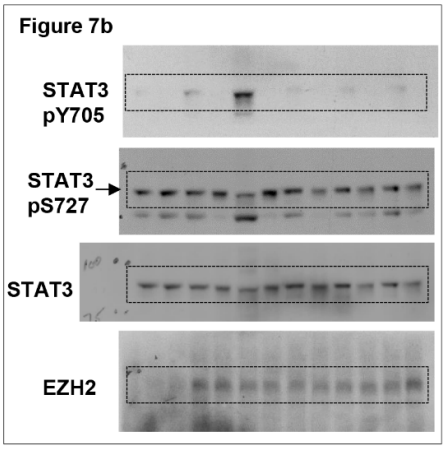


Figure 7b
STAT3 pY705
STAT3 pS727
STAT3
EZH2

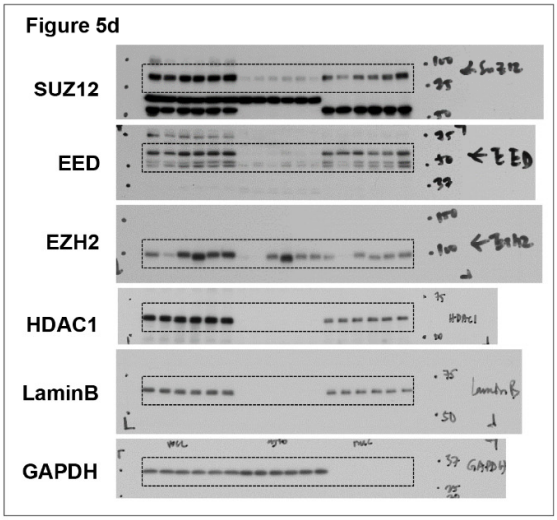


Figure 5d
SUZ12
EED
EZH2
HDAC1
LaminB
GAPDH

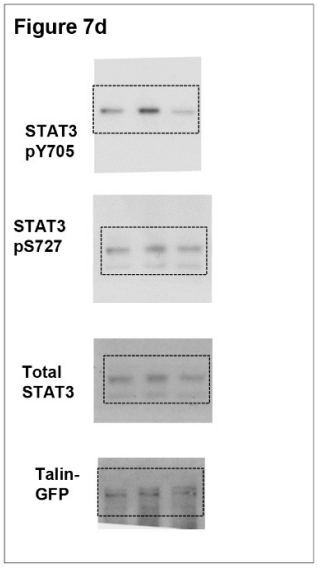


Figure 7d
STAT3 pY705
STAT3 pS727
Total STAT3
Talin-GFP

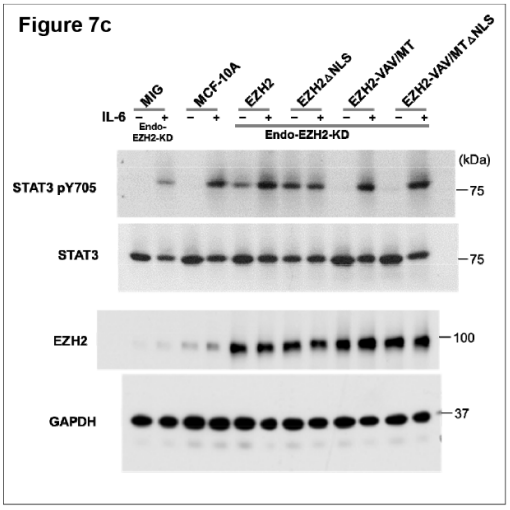


Figure 7c
MIG
MCF-10A
EZH2
EZH2ΔNLS
EZH2-VAV/MT
EZH2-VAV/MTΔNLS
IL-6
Endo-EZH2-KD
Endo-EZH2-KD
(kDa)
STAT3 pY705
75
STAT3
75
EZH2
100
GAPDH
37

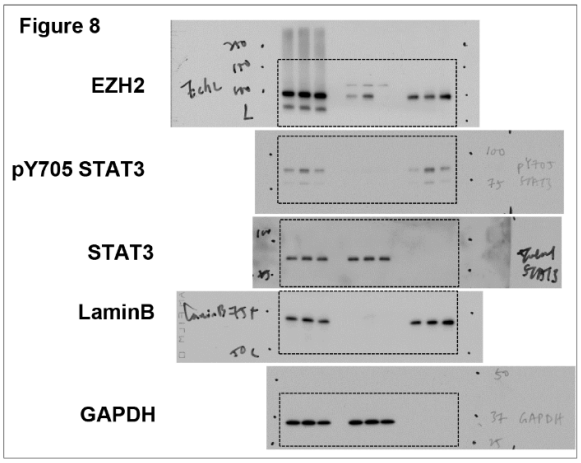
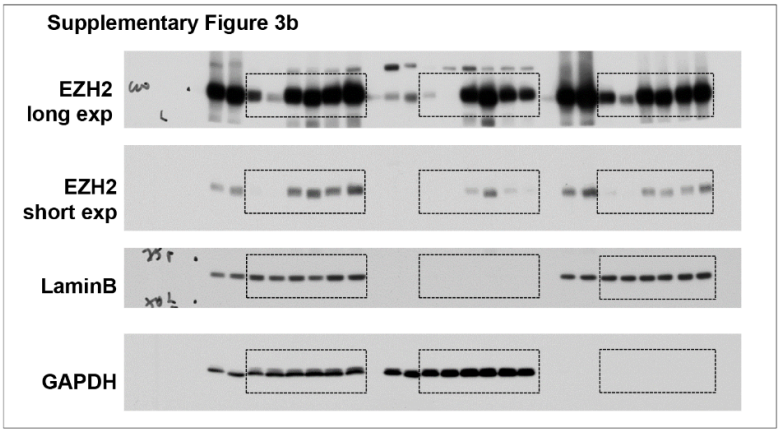
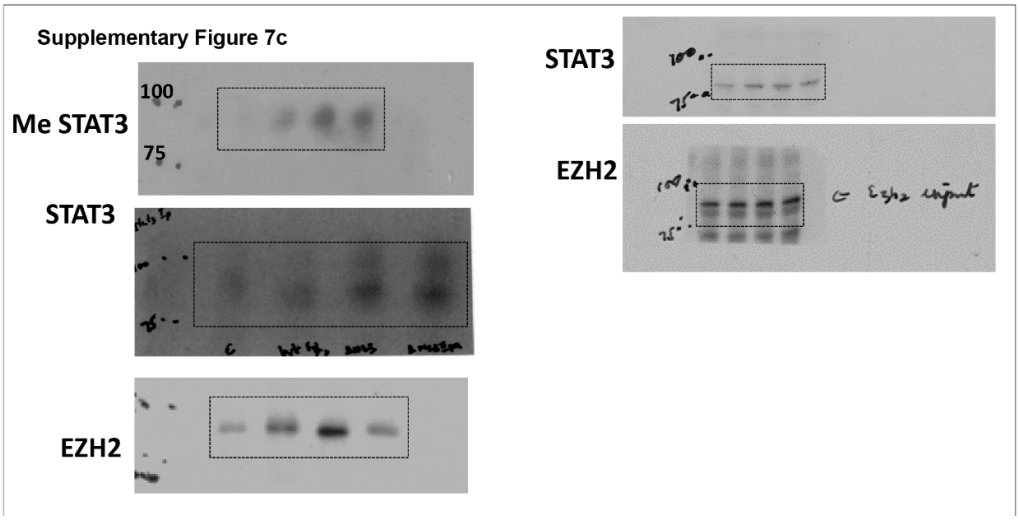


Figure 8
EZH2
pY705 STAT3
STAT3
LaminB
GAPDH


Supplementary Figure 3b
EZH2
long exp
EZH2
short exp
LaminB
GAPDH


Supplementary Figure 7c
Me STAT3
100
75
STAT3
EZH2
STAT3
EZH2

2. Supplementary Video Legends

Video 1 (related to Figure 4) EZH2ΔNLS promotes adhesion turnover.
The turnover of adhesion structures in EZH2 knockdown H16N2 cells overexpressing EZH2ΔNLS and mCherry-Paxillin was monitored by mCherry signals in time-lapse imaging.

Video 2 (related to Figure 4) VAV interaction mutant EZH2 prevents adhesion turnover. The turnover of adhesion structures in EZH2 knockdown H16N2 cells overexpressing VAV interaction mutant EZH2 (EZH2-VAV/MT) and mCherry-Paxillin was monitored by mCherry signals in time-lapse imaging.

Video 3 (related to Figure 5) Un-methylatable Talin1 contributes to extremely stable adhesion structures (Talin1-K2454Q).
The turnover of adhesion structures in unmethylatable GFP-Talin1 (Talin1-K2454Q) overexpressing H16N2 cells was monitored by GFP signals in time-lapse live imaging.

Video 4 (related to Figure 5). Methyl-mimicking Talin1 promotes adhesion turnover.
The turnover of adhesion structures in methyl-mimicking GFP-Talin1 (Talin1-K2454F) overexpressing H16N2 cells was monitored by GFP signals in time-lapse live imaging.

Video 5 (related to Figure 8) 3D reconstruction of Z-stack confocal images of EZH2 expression in breast cancer stem cells ($CD44^{hi}$SP).
Sorted $CD44^{hi}$ Rhodamine 123^{low} side population (SP) MDA-MB-231 cells were stained with anti-EZH2 specific antibody (green) and DAPI (blue) for nuclei.

Video 6 (related to Figure 8) 3D reconstruction of Z-stack confocal images of EZH2 expression in breast cancer cells ($CD44^{low}$).
Sorted $CD44^{low}$ Rhodamine 123^{hi} MDA-MB-231 cells were stained with anti-EZH2 specific antibody (green) and DAPI (blue) for nuclei.

3. Supplementary Methods

Structural Analysis of EZH2

Since the 52-residue sequence stretch of EZH2 that interacts with vav1 has no identifiably homologs of known structure, its 3D structure could not be accurately modeled. Searches of the domain databases PFam[1] and SMART[2] for this sequence stretch also did not reveal any known motifs. A secondary structure prediction of this part of the EZH2 sequence was hence done using GORIV[3], Jpred[4], PSI-pred[5] and HMM-STR[6]. A consensus of the predictions was taken.

Homology modeling of VAV1

Models of the 3D structure of the domain (residues 1 to 172) that contains the interacting stretch (residue 66 to 115) of VAV1 were built with MODELLER9.9 (Ref[7]) using its homolog VAV-3 (PDB:2d86; sequence identity of 71%) as structural template. The target-template alignment was done using SALIGN[8]. The models were assessed using the statistical potential DOPE[9]. APBS[10] was used to determine the surface electrostatics of the structure. For visualization and representation, ribbon diagrams and surface renditions of the proteins were done using UCSF Chimera[11]. DOPE evaluation of the VAV1 model: the VAV1 model has a normalized z-score of -1.469, which is comparable to its structural template VAV3 (-1.680). The target has a z-score that is less than -1, which is usually indicative of a native-like model.

Modeling the Interacting Interface

The 50-residue sequence stretch of VAV1 implicated in binding EZH2 has a calponin homology domain motif[12]. Using this sequence stretch, the PDB was searched for all known calponin homology domains using DBAli[13]. Templates to model the structure of the VAV1 – EZH2 interaction were chosen by manually filtering out non-interacting and homo-oligomeric calponin homology domains (see list below). The interaction templates hence consisted of 8 structures (PDB codes:1dxx, 1pxy (2 domains), 1qag, 1sh5, 1wku, 2r0o, 3hoc) where calponin homology domains were bound to another protein. A manual inspection of the template structures revealed that interactions between the calponin homology domain and its interactor were predominantly electrostatic in nature and mediated by a single helix. The calponin homology domain of VAV1 model was structurally aligned with the calponin homology domain of the 8 templates using CLICK[14,15]. This structural alignment of the calponin homology domains also structurally aligned with the 8 interacting helices (C^{α} RMSD between 0.31Å ~ 2.56Å). The interacting helices were not always part of the related protein structures and the sequence identities amongst them vary from 13% - 82%. The superimposed calponin homology domain and the helices they interact with served as the template to model the interaction between VAV1 and EZH2. To model the interacting part of EZH2, its secondary structure was predicted (see above). Helical regions were manually aligned to the interacting template helices. Using this alignment, a 3D model of the interface was built using MODELLER (version 9.9). The models of the interacting complex were assessed using the statistical potential[16].

Statistical potential evaluation of VAV1-EZH2 complex model

Five models of the VAV1 – EZH2 complex models were built. The average statistical interaction score[16] over the 5 models was -1.36 (standard deviation ~ 0.12), which is indicative of biologically meansingful interaction.

PDB files possibly containing Calponin Homology Domain

1bhd, 1dxx, 1pxy, 1qag, 1sh5, 1sjj, 1wku, 2bra, 2bry, 2c4c, 2k3s, 2qjz, 2r0o, 2ve7, 2vzc, 2vzd, 2vzg, 2vzi, 3hoc, 3hop, 3hor, 3i6x, 3kmu, 3kmw, 3ky9, 3rep, 1aa2, 1aca, 1aoa, 1bhd, 1bkr, 1bqb, 1dxx, 1ehj, 1h67, 1hmi, 1hmo, 1ho8, 1hy5, 1jaf, 1kaf, 1lm2, 1mb8, 1n91, 1nei, 1nwa, 1ogg, 1p2x, 1p5s, 1pa7, 1pow, 1pxy, 1rt8, 1s9g, 1sh5, 1sh6, 1t4o, 1t8k, 1th8, 1til, 1tjt, 1ueg, 1ufo, 1ujo, 1upt, 1uv7, 1v4s, 1v5k, 1vka, 1w1w, 1wix, 1wku, 1wud, 1wyl, 1wym,

1wyn, 1wyo, 1wyp, 1wyq, 1wyr, 1xhk, 1xvl, 1y7o, 1zt3, 1zt5, 2ahm, 2amw, 2b0h, 2bo4, 2bo8, 2c2l, 2cb8, 2cj7, 2cj8, 2csb, 2d0e, 2d85, 2d86, 2d87, 2d88, 2d89, 2da6, 2dk9, 2dsq, 2dym, 2dyo, 2eyi, 2eyn, 2ez4, 2ez8, 2ez9, 2ezt, 2ezu, 2f5z, 2flf, 2fq0, 2fq2, 2fr0, 2fr1, 2fzv, 2mhr, 2new, 2o1e, 2ok7, 2oo2, 2ooc, 2ot8, 2ov1, 2p32, 2ps3, 2q4f, 3ygs

RNA-seq

H16N2 human normal epithelial cells (2×10^5) were seeded in 6-well plates for 48 hours. RNA was isolated using QIAshredder (Qiagen) and RNeasy kit (Qiagen). The total RNA quality and RIN value was confirmed with a 2100 Bioanalyzer (Agilent) using RNA 6000 Nano chip (Agilent). Library preparation was done using the Illumina® TruSeq Stranded mRNA LT Sample Prep Kit, following the manufacturer's protocol. During library construction, enriched mRNA was fragmented to obtain inserts with fragment size of 120–200bp. Each sample was linearized, denatured and loaded into 2 lanes of Single-Read flowcell using the Illumina cBOT. The cDNA were attached to the flowcell surfaces and amplified to clusters and attached with sequencing primers. The flowcell was then loaded into the Illumina HiSeq2000 Sequencers with the Illumina® HiSeq SBS Kit and run at 1 x 101 cycles, generating Single-Read 100 base-pair reads. The images were captured by the HiSeq Control Software (HCS), and the Real Time Analysis (RTA) software converted the images into Cycle Intensity Files (CIF) and subsequently Basecall (bcl) files. RNA-Seq data were mapped against the human genome version hg19 with TopHat2[17], using the GENCODE Release 19 version of gene annotations. FPKM values were extracted using Cufflinks 2.2.1 (Ref[18]). R-3.1.2 (Ref[19]) and Bioconductor 3.0 (Ref[20,21]) were used for the RNA-Seq analysis. Reads were counted using the R package GenomicAlignments[22] (mode='Union', inter.feature=FALSE). Only primary read alignments were retained. Differentially expressed genes were identified using DESeq2 (Ref[23]). Differentially expressed genes for the 4 conditions (MIG, MIG-EZH2ΔNLS, MIG-EZH2H689A, and MIG-EZH2-VAVMT) were calculated in comparison to MIG-EZH2. If the log2-fold change of a gene was higher than 0 and its adjusted p-value was smaller than 0.05 the gene was highlighted as upregulated. If the log2-fold change of a gene was smaller than 0 and its adjusted p-value was smaller than 0.05, the gene was highlighted as downregulated. The heatmap was created by using ggplot2_1.0.0 (Ref[24]). 2176 significantly up or down regulated genes were found in the MIG-EZH2 and MIG-EZH2ΔNLS comparison. Normalized FPKM values of these significantly expressed genes were shown as a heatmap in **Supplementary Figure 7a**. The data was normalized such that the sum of the expression estimates of each gene was 1 (relative expression) and the normalized FPKM values from the EZH2 sample were subtracted from the other samples in order to show the EZH2 sample as the baseline.

REFERENCES:

1. Bonneau, R. et al. De novo prediction of three-dimensional structures for major protein families. *J Mol Biol* **322**, 65-78 (2002).
2. Schultz, J., Milpetz, F., Bork, P. & Ponting, C.P. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-64 (1998).
3. Garnier, J., Gibrat, J.F. & Robson, B. GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol* **266**, 540-53 (1996).
4. Cuff, J.A. & Barton, G.J. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins* **40**, 502-11 (2000).
5. McGuffin, L.J., Bryson, K. & Jones, D.T. The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-5 (2000).
6. Bystroff, C., Thorsson, V. & Baker, D. HMMSTR: a hidden Markov model for local sequence-structure correlations in proteins. *J Mol Biol* **301**, 173-90 (2000).

7. Sali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815 (1993).
8. Madhusudhan, M.S., Marti-Renom, M.A., Sanchez, R. & Sali, A. Variable gap penalty for protein sequence-structure alignment. *Protein Eng Des Sel* **19**, 129-33 (2006).
9. Shen, M.Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci* **15**, 2507-24 (2006).
10. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. & McCammon, J.A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-41 (2001).
11. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
12. Castresana, J. & Saraste, M. Does Vav bind to F-actin through a CH domain? *FEBS Lett* **374**, 149-51 (1995).
13. Marti-Renom, M.A., Ilyin, V.A. & Sali, A. DBAli: a database of protein structure alignments. *Bioinformatics* **17**, 746-7 (2001).
14. Nguyen, M.N. & Madhusudhan, M.S. Biological insights from topology independent comparison of protein 3D structures. *Nucleic Acids Res* **39**, e94 (2011).
15. Nguyen, M.N., Tan, K.P. & Madhusudhan, M.S. CLICK--topology-independent comparison of biomolecular 3D structures. *Nucleic Acids Res* **39**, W24-8 (2011).
16. Davis, F.P. et al. Protein complex compositions predicted by structural similarity. *Nucleic Acids Res* **34**, 2943-52 (2006).
17. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36 (2013).
18. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-5 (2010).
19. Team, R.D.C., R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2014).
20. Huber, W., et al., Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **12**(2), 115-121 (2015).
21. Gentleman, R.C., et al., Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**(10), R80 (2004).
22. Lawrence, M., et al., Software for computing and annotating genomic ranges. *PLoS Comput Biol* 9(8), e1003118 (2013).
23. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**(12), 550 (2014).
24. Wickham, H., ggplot2: elegant graphics for data analysis, Springer New York (2009).