Biochimica et Biophysica Acta 1849 (2015) 731–742

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Enhancer cooperativity as a novel mechanism underlying the transcriptional regulation of E-cadherin during mesenchymal to epithelial transition

Hani Alotaibi [a,b], M. Felicia Basilicata [a], Huma Shehwana [c], Tyler Kosowan [a], Ilona Schreck [a], Christien Braeutigam [a], Ozlen Konu [c], Thomas Brabletz [d,e,f,g], Marc P. Stemmler [a,d,g,*]

[a] Department of Molecular Embryology, Max-Planck Institute of Immunobiology and Epigenetics, Stuebeweg 51, D-79108 Freiburg, Germany
[b] Izmir Biomedicine and Genome Center, Dokuz Eylül University, Inciralti, 35340 Izmir, Turkey
[c] Department of Molecular Biology and Genetics, Bilkent University, 06800 Ankara, Turkey
[d] Department of Visceral Surgery, University Medical Center Freiburg, Hugstetter Str. 55, D-79106 Freiburg, Germany
[e] Comprehensive Cancer Center Freiburg, University Medical Center Freiburg, D-79106 Freiburg, Germany
[f] BIOSS Centre for Biological Signaling Studies, Albert-Ludwigs-University Freiburg, D-79104 Freiburg, Germany
[g] Institute of Experimental Medicine I, Nikolaus-Fiebiger-Center for Molecular Medicine, University Erlangen-Nürnberg, D-91054 Erlangen, Germany

ARTICLE INFO

Article history:
Received 13 October 2014
Received in revised form 6 January 2015
Accepted 24 January 2015
Available online 31 January 2015

Keywords:
Cadherins
Grhl3
Hnf4α
Transforming growth factor beta
Development
Cancer

ABSTRACT

Epithelial–mesenchymal transition (EMT) and mesenchymal–epithelial transition (MET) highlight crucial steps during embryogenesis and tumorigenesis. Induction of dramatic changes in gene expression and cell features is reflected by modulation of *Cdh1* (E-cadherin) expression. We show that *Cdh1* activity during MET is governed by two enhancers at +7.8 kb and at +11.5 kb within intron 2 that are activated by binding of Grhl3 and Hnf4α, respectively. Recruitment of Grhl3 and Hnf4α to the enhancers is crucial for activating *Cdh1* and accomplishing MET in non-tumorigenic mouse mammary gland cells (NMuMG). Moreover, the two enhancers cooperate via Grhl3 and Hnf4α binding, induction of DNA-looping and clustering at the promoter to orchestrate E-cadherin re-expression. Our results provide novel insights into the cellular mechanisms whereby cells respond to MET signals and re-establish an epithelial phenotype by enhancer cooperativity. A general importance of our findings including MET-mediated colonization of metastasizing tumor cells is suggested.

1. Introduction

Many key steps during embryogenesis result in the formation of new cell types with unique features. They become morphologically visible when individual cells or tissues are generated by cell delamination during a process called epithelial–mesenchymal transition (EMT) or by cell clustering and re-epithelialization during mesenchymal–epithelial transition (MET). EMT is required for mesoderm formation, neural crest cell delamination, fibrosis and wound healing, but is also aberrantly activated during tumorigenesis when cancer cells start to disseminate, invade and form metastases [1–3]. Common to all types of EMT are cytoskeletal rearrangements resulting in loss of cell polarity and adherent morphology combined with increased migration. Gene expression signatures are changing dramatically with a major impact on the repertoire of cell adhesion molecules, especially of Ca^{2+}-dependent adhesion molecules, the cadherins [4]. Strikingly, changes in cellular characteristics during a bona fide EMT are to a large extent dependent on the downregulation of E-cadherin (E-cad) and the activation of N-cadherin (N-cad), regulated by the EMT program [1].

MET is considered as the reverse process of EMT and also originates from embryogenesis [5,6]. Here, mesenchymal cells acquire epithelial characteristics including loss of N-cad and activation of E-cad expression [1–3]. In addition to orchestrating morphogenetic events during embryogenesis the process of MET is utilized by disseminating tumor cells and is required for colonization and formation of metastases at distant sites [1,2,7]. In vitro MET is necessary for somatic cell reprogramming. During the generation of induced pluripotent stem (iPS) cells, MET precedes the activation of the endogenous loci of transcription factors of the core pluripotency network [8]. This is in part established by the activation of E-cad expression via direct binding of exogenous Klf4 to specific sites at the promoter [9]. Interestingly, E-cad supports the initiation of MET and increases reprogramming efficiency [10]. In contrast, E-cad depletion results in loss of pluripotency and decreases the potential for reprogramming [11]. All these findings depict EMT, MET and a tightly regulated expression of cadherins as crucial during many different processes in development and disease.

* Corresponding author. Tel.: +49 9131 85 29101; fax: +49 9131 85 29341.
E-mail address: marc.stemmler@fau.de (M.P. Stemmler).

http://dx.doi.org/10.1016/j.bbagrm.2015.01.005
1874-9399/

Hallmarks of transitions between epithelial and mesenchymal states include changes in the cadherin repertoire. An essential step of EMT is the downregulation of E-cad whereas its expression is rapidly restored during MET. Many different extracellular signals are known to trigger EMT leading to activation of intracellular EMT-inducers like Snail, Slug, Twist, Zeb1, Zeb2 and others [1,2]. In agreement with a required fast downregulation of E-cad, all of these transcription factors are in fact repressors of E-cad expression and they all bind to several evolutionary conserved E-boxes present in the proximal promoter [12–18]. However, the immediate activation of the *Cdh1* locus during MET can only in part be explained by the downregulation and release of these repressors. We found that the *Cdh1* promoter alone, including all known E-box elements, is insufficient to confer strict cell type specificity. In contrast, intron 2 carries sufficient information for proper E-cad expression [19,20]. In particular, we identified sequences in the proximal 15 kb of the 45 kb spanning intron 2, that showed cell type specific gene activation, mainly in the endoderm [20]. Moreover, an endogenous *Cdh1* allele lacking the entire intron 2 failed to activate expression of a reporter, showing that intron 2 is essential for the initiation and maintenance of E-cad expression during development [19]. Although intron 2 has emerged as an essential regulator of *Cdh1* expression, very little is known about the molecular determinants controlling the transcriptional activity of the locus. Besides Klf4, it has been shown that Grainyhead-like 2 (Grhl2), a mammalian homolog of *Drosophila* grainyhead, controls epithelial differentiation in several epithelia and at uretic bud formation by regulating E-cad via binding to an element in intron 2 [21].

Here, we sought to identify novel enhancers within intron 2 of *Cdh1* and to discover the corresponding transcription factors that mediate gene activation. We used non-tumorigenic mouse mammary gland epithelial cells (NMuMG) that undergo a reversible EMT by transforming growth factor β (TGFβ) treatment. We identified two potent enhancers that cooperate in establishing prompt and robust E-cad expression during the induction of MET. Furthermore, we provide evidence that transcriptional activity from these enhancers is mediated by two novel E-cad activators, Grhl3 and Hnf4α, which emerge as regulators of MET in NMuMG cells.

2. Materials and methods

2.1. Cell culture and induction of EMT and MET

NMuMG; non-tumorigenic mouse mammary gland cell line, Hepa1–6; mouse hepatoma cell line, CMT; mouse polyploid carcinoma cell line and HEK-293; human embryonic kidney cell line were obtained from ATCC. Cells were maintained in DMEM (Gibco) and 10% fetal bovine serum (FBS) at 37 °C, 10% CO_2. NMuMG medium was supplemented with 10 μg/ml insulin (Sigma). EMT induction of NMuMG cells was done by applying 5 ng/ml TGFβ3 (PeproTech) for 72 h. The reversible transition to the epithelial state (MET) was induced by TGFβ3 withdrawal, washing the plates twice with PBS and incubation for additional 72 h in fresh medium.

2.2. Luciferase reporter assays

Reporter plasmids and expression vectors used here are available in the Supplemental Experimental Procedures. Cells in 48-well plates were transfected with plasmid DNAs using X-tremeGENE 9 transfection reagent (Roche). Transfection was carried out with 100 ng DNA containing 5 ng pRL-TK (Promega) to normalize for transfection efficiency. We normalized to equal molarity of the plasmid to use equal copies of reporter plasmid. Absolute DNA content was equalized by addition of promoter-less plasmid DNA. In experiments with TGFβ3, the drug was added at the time point of transfection. Cells recovering from TGFβ3 treatment (PT) were transfected upon TGFβ3 withdrawal and measured after 3 days of recovery. 20 ng of *Grhl3*, *Grhl2*, *Grhl1* and *Hnf4a* expression plasmid or empty vector (mock) were transfected. Luciferase reporter activity was measured using Dual-Glo Luciferase Assay (Promega) in a Centro LB 960 luminometer (Berthold Technologies). Firefly luciferase reporter values were normalized to those of the Renilla luciferase control. Fold induction was calculated relative to the values of the empty vector or vehicle control samples.

2.3. Generation of transgenic embryos

Animal husbandry and all experiments were performed according to the German Animal Welfare guidelines and approved by the local authorities. The reporter construct TG3 (Wt) or the one with mutated Grhl3 binding site 7.8b (Mut) were injected into oocytes as described [20].

2.4. Chromosome conformation capture (3C)

The protocol for 3C was described previously [22]. In brief, cells were grown in the presence or absence of 5 ng/ml TGFβ3 and allowed to recover after TGFβ3 withdrawal and cross-linked with 1% formaldehyde for 10 min at room temperature. After lysis on ice for 10 min in lysis buffer (10 mM Tris–HCl, pH 7.5; 10 mM NaCl; 5 mM $MgCl_2$; 0.1 mM EGTA; 1× complete protease inhibitor), nuclei were incubated with *Taq*I for 90 min at 37 °C. Purified DNA was used together with BAC RP23-262 N14 DNA containing the murine *Cdh1* locus for qPCR analysis. Primers used are presented in Supplementary Table S4.

2.5. Chromatin immunoprecipitation (ChIP)

ChIP was performed as reported previously [23] with minor modifications as depicted in the Supplemental Experimental Procedures using Grhl3 (S-19, Santa Cruz), Hnf4α (H-171, Santa Cruz), p300 (C-20, Santa Cruz), H3K9Ac (ab4441, Abcam), H3K4me1 (ab8895, Abcam) and a rabbit control IgG (sc-2027, Santa Cruz) antibodies.

2.6. Co-immunoprecipitation and immunoblotting

Immunoprecipitation was done as described in detail in the Supplemental Experimental Procedures using 1 mg nuclear extracts and 25 μl of anti-HA affinity matrix (Roche) or 4 μg of anti-FLAG antibody (M2, Sigma) coupled to Protein-G Dynabeads.

2.7. Statistical analysis

Statistical significance was determined by performing the Student's *t*-test using the normalized values of each test sample compared to the normalized value of the control, using a 95% confidence interval; p-values less than 0.05 were considered significant. Data are presented as the mean of at least three independent experiments, done in triplicates. ChIP and 3C were carried out three times, qPCR was performed in triplicates. Error bars represent standard error of the mean.

3. Results

3.1. Intron 2 of Cdh1 harbors conserved and potent enhancers

We reported previously that intron 2 of *Cdh1* is essential for proper E-cad expression [19,20]. In order to analyze this large DNA sequence (>40 kb) for putative enhancers we focused on evolutionary conserved sequences. We rationalized that a conserved expression pattern is reflected by a conserved regulatory mechanism utilizing the same enhancers conserved in highly similar genomic regions among different species [24]. To identify such putative enhancers, we used the mVista

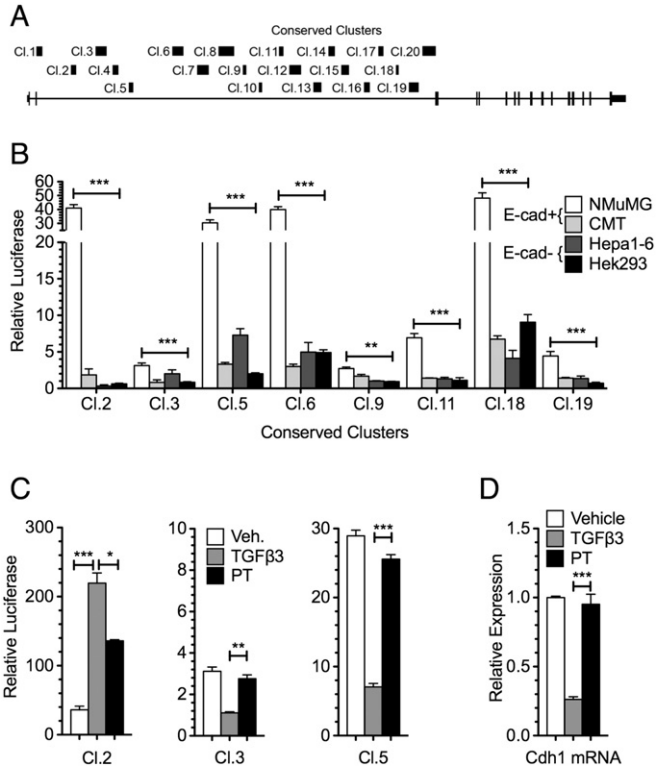

Fig. 1. Identification of functional conserved sequences within *Cdh1* intron 2. (A) A map of the conserved clusters is shown in relation to the *Cdh1* gene, location and size of clusters are in scale. Cl.1–Cl.20 indicate individual clusters and refer to the Luciferase constructs below (Supplementary Table S1). (B) Luciferase assay of constructs harboring conserved clusters in E-cad expressing and E-cad negative cell lines. Relative luciferase is calculated as fold induction relative to the empty vector pGL4.23. (C) Reporter gene activity of selected clusters in NMuMG cells in response to three days of vehicle and TGFβ3 treatment as well as after three days of TGFβ3 withdrawal (post-treatment, PT). (D) qRT-PCR analysis of endogenous E-cad mRNA in response to TGFβ3. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

multiple alignment tool [25] for cross-species sequence comparison of the *Cdh1* locus, and compared the genome of mouse to those of human, rhesus, dog, and horse (Supplementary Fig. S1A, Supplementary Table S1). We identified several sequences with more than 70% conservation in a 150 bp window consistent with our previous analysis [19]. These conserved clusters were distributed over the entire intron 2 (Fig. 1A). Individual clusters were cloned into a luciferase reporter vector, upstream of a minimal promoter (Supplementary Fig. S1B) and their enhancing potential was analyzed by measuring luciferase activity. We used two E-cad expressing (non-tumorigenic mammary gland and rectal polyploid carcinoma cells, NMuMG and CMT, respectively) and two E-cad negative cell lines (immortalized embryonic kidney and hepatoma cells, HEK-293 and Hepa1–6). To identify the most important sequences, the clusters were sorted out with respect to importance for E-cad regulation by two criteria. First we selected conserved sequences with stronger activity in cells expressing E-cad compared to those which are E-cad negative (Fig. 1B). The analysis showed that 8 out of the 20 clusters had stronger activity in the E-cad + cell line NMuMG when compared to the E-cad− cell lines (Fig. 1B, Supplementary Fig. S1B). None of the clusters showed strong activity in CMT cells, indicating a more complex mode of regulation exceeding the limitation of this technique (Fig. 1B, Supplementary Fig. S1B). As a second criterion 8 clusters with a higher activity in E-cad + cell lines were tested for changes in reporter activity during EMT in response to TGFβ3 treatment and during MET after TGFβ3 withdrawal (post-treatment, PT) in NMuMG cells (Fig. 1C, Supplementary Fig. S1C). In total three conserved sequences were found to be responsive to TGFβ3, but the reporter activity of only two (Cl.3 and Cl.5) recapitulated the expression of E-cad in response to TGFβ3, whereas Cl.2 showed inverse correlation (Fig. 1C and D).

3.2. Grhl3 but neither Grhl1 nor Grhl2 activates Cdh1 via an intronic enhancer in Cl.3 in NMuMG cells

We used the MatInspector tool from the Genomatix Suite [26] to screen Cl.3 and Cl.5 for transcription factor binding sites. We identified a recognition motif in Cl.3 for grainyhead-like transcription factors, proteins known to be required for pattern specification and tissue development in several species including Drosophila [27–30]. In mice it has been recently shown that Grhl2 is controlling E-cad expression during MET of uretic bud formation by binding to two highly conserved Grhl motifs in tandem within intron 2 located at 7.8 kb (7.8a and 7.8b) relative to the transcription start site (TSS) [21] (Supplementary Fig. S2A). We searched for presence of Grhl binding sites in intron 2 and restricted the analysis to the proximal 11 kb, which we previously identified as being sufficient to faithfully recapitulate endodermal E-cad expression in transgenic reporter mice [20]. In total we identified 17 putative Grhl binding sites, five of which are located within the 11 kb fragment (Fig. 2A, Supplementary Table S2). We tested whether the recognition motifs were also functional Grhl2 binding sites and utilized for *Cdh1* regulation in NMuMG cells by luciferase reporter assays. The architecture of the locus was maintained in the reporter constructs by cloning the luciferase reporter between the promoter (nucleotides from −1490 to +1) and the entire region between +0.1 kb and +11 kb of the *Cdh1* gene (Cl.1–4) (Fig. 2A). To our surprise, we found that Cl.1–4 was not activated by neither *Grhl2* nor *Grhl1*. In contrast, ectopic *Grhl3* expression increased reporter gene activity 4-fold (Fig. 2B, Supplementary Fig. S2F).

We also created several deletion constructs for analysis of individual binding sites that contained one or two of the four conserved clusters from the analysis in Fig. 1. In addition to the original 11 kb fragment, containing Cl.1–4 and sites 1.4, 2.4, 7.8a, 7.8b and 9.6, only deletion constructs containing Cl.3 were significantly activated by Grhl3 (Fig. 2B). Including downstream sequences up to 16 kb containing Cl.1–5 and sites 10.7, 12.7 and 13.7 did not significantly increase reporter gene activity (Supplementary Fig. S2B and C). Furthermore, we introduced point mutations in the Grhl3 binding sites at 7.8a, 7.8b and 9.6 (Supplementary Fig. S2D), and found that the mutation at site 7.8b already reduced the basal activity, indicating that endogenous Grhl3 is no longer able to bind and activate the construct. Moreover, this mutation completely abolished the Grhl3-dependent upregulation (Fig. 2C) and resulted in impaired regulation of the reporter during TGFβ3 treatment (Fig. 2D). In contrast, mutating site 7.8a had only moderate effects and the construct with a mutation at site 9.7 retained full activity in response to Grhl3 (Fig. 2C, Supplementary Fig. S2C). This suggested a possible role for Grhl3 in regulating transcription from the enhancer at site 7.8b in response to TGFβ3.

In order to assess the functionality of this enhancer and the requirement for Grhl3-dependent regulation in vivo independent of MET events, we used the mouse transgenic *lacZ*-reporter TG3, corresponding to Cl.1–4 [20]. We introduced a point mutation in the Grhl3 binding site at 7.8b and analyzed the reporter expression in the developing mouse embryo at E11.5. In agreement with our previous findings, embryos injected with the wildtype reporter (Wt, n = 9/9 X-gal positive embryos) were characterized by β-gal expression in the endoderm of the pharynx, esophagus, lung, stomach and pancreas, reflecting an expression pattern similar to that of E-cad. To our surprise, this expression was absent in embryos carrying the mutant construct (Mut, n = 0/16 X-gal positive embryos), indicating an important role for the Grhl3 motif in regulating the expression of E-cad in endodermal epithelia of the developing mouse embryo (Fig. 2E, and Supplementary Fig. S2E).

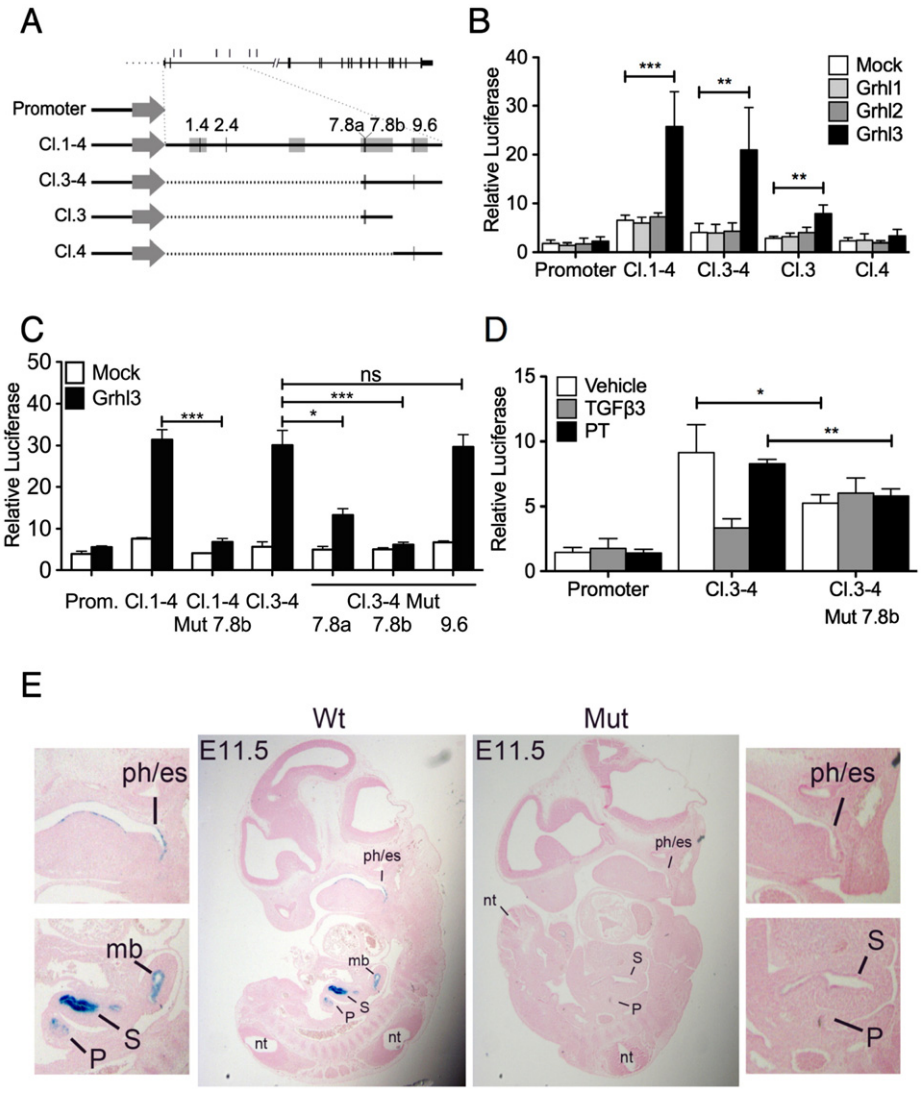

Fig. 2. Grhl3 activates a potent enhancer in Cl.3 of intron 2. (A) A schematic representation of the *Cdh1* locus is given in the upper panel, with the vertical lines above the locus map depicting the identified Grhl-binding sites between −1.5 and +16 kb. Below is a representation of the maps of the luciferase reporter vectors under the control of the *Cdh1* promoter. Shaded boxes represent conserved clusters included in the constructs. Numbers and vertical lines represent relative positions of Grhl binding motifs and dashed lines indicate excluded sequences (Supplementary Table S2). (B) Luciferase assays in NMuMG cells showing the effect of Grhl1, Grhl2 or Grhl3 on the indicated reporter construct. (C) Luciferase assays in NMuMG cells studying the Grhl3 function in response to point mutations in the Grhl-binding sites at 7.8a + b and 9.6. (D) Analysis of the construct with the mutated Grhl-binding site after TGFβ3 treatment and after withdrawal. Relative luciferase activity is calculated as fold induction relative to vector (B and C) or to vehicle (D; Supplementary Table S3). (E) Analysis of transgenic embryos using a control *lacZ*-reporter construct similar to Cl.1–4 (Wt) and the corresponding one that harbors a mutation in the Grhl-binding site at +7.8 (Mut). Representative midsagittal sections of X-gal stained E11.5 embryos are shown with higher magnifications of the X-gal positive domains.

3.3. Grhl3 binding to E-cad chromatin occurs during mesenchymal to epithelial transition

When we investigated the presence of Grhl3 on the *Cdh1* chromatin in untreated NMuMG cells by ChIP as an indication of direct binding, we could not detect significant enrichment of Grhl3 on any of the predicted sites, including the functional site at 7.8b under steady state conditions (Fig. 3A, vehicle). We then took advantage of the ability of these cells to undergo TGFβ-dependent reversible EMT. TGFβ3 treatment for 72 h induced EMT, indicated by downregulation of E-cad and upregulation of Vimentin. Upon withdrawal of TGFβ3 the cells undergo MET resulting in re-expression of E-cad (Supplementary Fig. S3A and B). We investigated a possible enrichment of Grhl3 on E-cad during MET after TGFβ3 withdrawal. We found that Grhl3 was significantly enriched at two of the Grhl3 sites (sites 2.4 and 7.8a + b) and also at the TSS 72 h post-treatment (Fig. 3A, PT), but not in cells treated with vehicle control or TGFβ3 (Fig. 3A).

In an effort to mechanistically address the observed Grhl3-dependent regulation of E-cad during MET, we investigated the enhancer–promoter interaction. We designed 3C experiments and compared the relative crosslinking frequencies in NMuMG cells in response to EMT–MET stimuli. Changes in chromosome conformation (DNA looping) can be depicted by measuring crosslinking frequencies of intermolecular ligation products, which represent enhancer–promoter interactions [22]. We used several primers located within intron 2 (Fig. 3B, Supplementary Table S4) and performed the qPCR analysis using a common anchoring primer near the promoter (RT1). We detected elevated crosslinking frequencies with primers corresponding to the location of the enhancers at 2.4 and 7.8a + b (FI2 and FG2; fragments 3 and 6 respectively) compared to primers located elsewhere between +1 and +16 kb (Fig. 3B). Importantly, the crosslinking frequencies of cells undergoing MET were significantly higher than of those treated with TGFβ3 or vehicle, suggesting that such enhancers were in close proximity to the promoter. These results are supporting the ChIP experiments, suggesting that in the initial

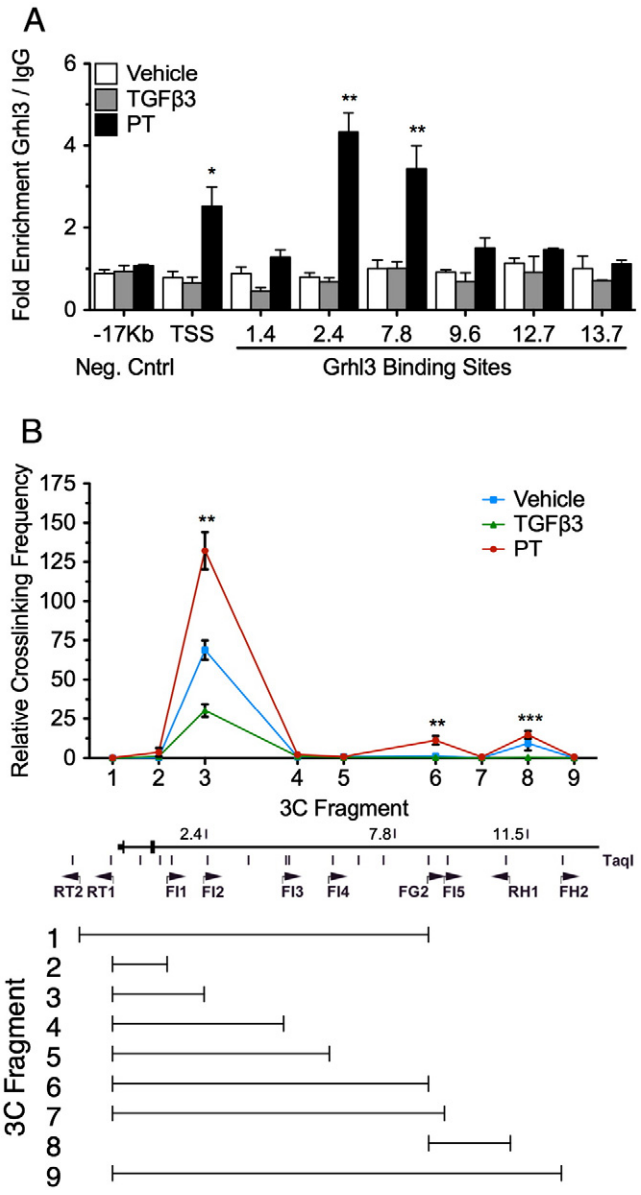

Fig. 3. Grhl3 is enriched at the intronic enhancers during mesenchymal–epithelial transition (MET). (A) ChIP with anti-Grhl3 antibodies in NMuMG cells undergoing EMT-MET in response to TGFβ3 treatment. Analysis of Grhl-binding sites in comparison to a non-related control sequence (−17 kb) shows specific enrichment at the Cl.3 enhancer containing binding site +7.8 and at a novel site at +2.4 as well as at the transcriptional start site (TSS; Supplementary Table S5). (B) Chromosome conformation capture (3C) experiment of the *Cdh1* locus in response to TGFβ3. The schematic representation of the *Cdh1* locus indicating the location of the enhancers (vertical lines above) the location of TaqI restriction sites (vertical lines below), and the location of the primers used (arrow heads represent orientation). The capped lines below indicate the primer pair used to amplify each 3C fragment (Supplementary Table S4 and S5).

phase after TGFβ3 withdrawal, Grhl3 is recruited to the enhancer. This may lead to an interaction of the Grhl3–DNA enhancer complex with the basal transcription machinery assembled at the promoter. Alternatively, binding of a preassembled Grhl3/basal transcription machinery/co-factor complex is stabilized by Grhl3–DNA binding at the enhancer. Both scenarios are in line with a gradual increase in *Cdh1* mRNA levels after TGFβ3 withdrawal (Fig. 1D). No evidence for Grhl3 binding to the putative site at 2.4 was found by luciferase reporter assays in NMuMG cells (Supplementary Fig. S2C). A mutation at site 7.8b alone was sufficient to abrogate Grhl3-dependent luciferase reporter activation and also the *lacZ*-reporter activity in transgenic embryos. In contrast, 3C and ChIP data revealed that this enhancer is utilized and recruits Grhl3 to the site at 2.4 in a context-dependent manner. Specifically, it is active during MET after TGFβ3 withdrawal. Some of the primers used in this experiment were in close proximity to the conserved sequence Cl.5 from Fig. 1 (primers FH2 and RH1; fragment 9). Although there was no detectable interaction between this conserved sequence and the promoter in this context, we were able to detect a significant interaction with the enhancer at 7.8a + b (primers RH1 and FG2; fragment 8), which was also a TGFβ3 responsive interaction detectable during MET (Fig. 3B, PT), suggesting a possible interplay between these two enhancers at 11.5 and 7.8a + b.

We then analyzed the effect of Grhl3 downregulation on E-cad expression in NMuMG cells using several *Grhl3* siRNAs. We found that two of the three analyzed siRNAs resulted in more than 80% downregulation of *Grhl3* (Fig. 4A). This reduction in *Grhl3* expression resulted in a significant decrease in E-cad expression levels only in cells recovering from TGFβ3 treatment (Fig. 4A), but not in cells treated with the vehicle (Supplementary Fig. S3C). We also noticed that with increasing

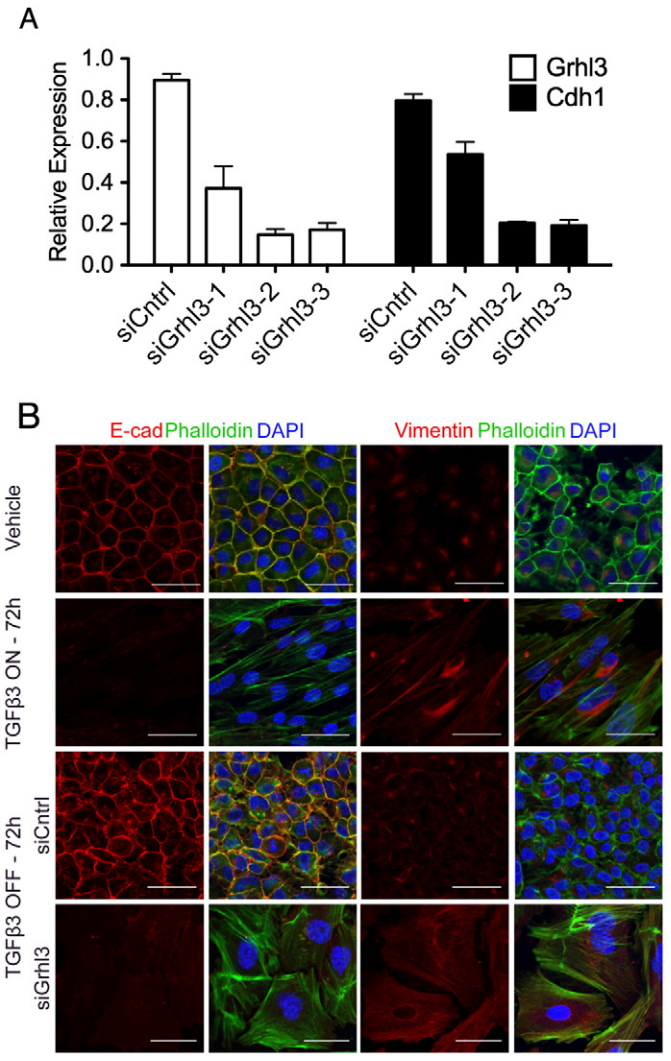

Fig. 4. Grhl3 depletion results in E-cad downregulation and a failure of MET. (A) qPCR analysis of mRNA of NMuMG cells recovering from TGFβ3 treatment after Grhl3 knockdown using three different siRNAs and the effect on E-cad mRNA levels (Supplementary Table S6). (B) Immunofluorescence labeling and phalloidin staining of NMuMG cells visualizing the change in expression and intracellular distribution of E-cad, Vimentin and Actin during TGFβ3 treatment and withdrawal. Actin distribution (Phal, green) and detection of E-cad (red) and Vimentin (red) is shown. Nuclei are labeled with DAPI. Scale bar, 50 μm.

efficiencies of *Grhl3* silencing, the effect on E-cad downregulation was more apparent (Fig. 4A, compare siGrhl3-1 and siGrhl3-2). To our surprise, the siRNA mediated loss of *Grhl3* expression after TGFβ3 withdrawal blocked MET (Fig. 4B). This was evident by the absence of E-cad expression, elevated Vimentin expression and the arrangement of cortical actin in stress fibers, reminiscent of continuous TGFβ3-mediated EMT (Fig. 4B, Supplementary Fig. S3D and E). A similar block of MET can also be observed in cells transfected with siRNAs targeting E-cad during the recovery from TGFβ3 treatment (Supplementary Fig. S3E). This indicated that Grhl3 function in this context is indispensable for proper MET. In part this is accomplished by Grhl3 positively regulating the expression of E-cad, which is essential for proper transition to the epithelial state after TGFβ3 withdrawal.

3.4. Hnf4α specific regulation of E-cad via Cl.5 in a TGFβ-dependent manner

Our analysis showed that the mode of action of the enhancer within Cl.3 is mainly mediated by Grhl3 during MET of NMuMG cells and during embryogenesis in the endoderm. Next, we investigated the contribution of Cl.5 to *Cdh1* gene regulation in mammary epithelial cells before and after TGFβ3 treatment. We searched the sequences of Cl.5 for the presence of putative transcription factor binding sites using the MatInspector tool of the Genomatix suite [26]. Among the several hits obtained we were particularly interested in a highly conserved recognition motif for the orphan nuclear receptor hepatocyte nuclear factor 4 alpha (Hnf4α) located at 11.5 relative to the TSS (Fig. 5A, Supplementary Fig. S1A). Hnf4α, which is faithfully coexpressed with E-cad, was previously shown to regulate several adhesion molecules (including E-cad) during fetal liver organogenesis [31]. Using luciferase reporter assays, we first tested whether Hnf4α could activate Cl.5 in NMuMG cells. We found that indeed Hnf4α was able to exert a robust enhancing effect on the E-cad Cl.5 reporter construct. Two different mutations in the putative Hnf4α binding site decreased the basal activity of this construct and inhibited its Hnf4α-dependent activation (Fig. 5B). Furthermore, we found that the two mutations in the Hnf4α consensus motif were sufficient to abolish recovery of the reporter activity upon withdrawal of TGFβ3 (Fig. 5C), suggesting that this recovery is Hnf4α dependent.

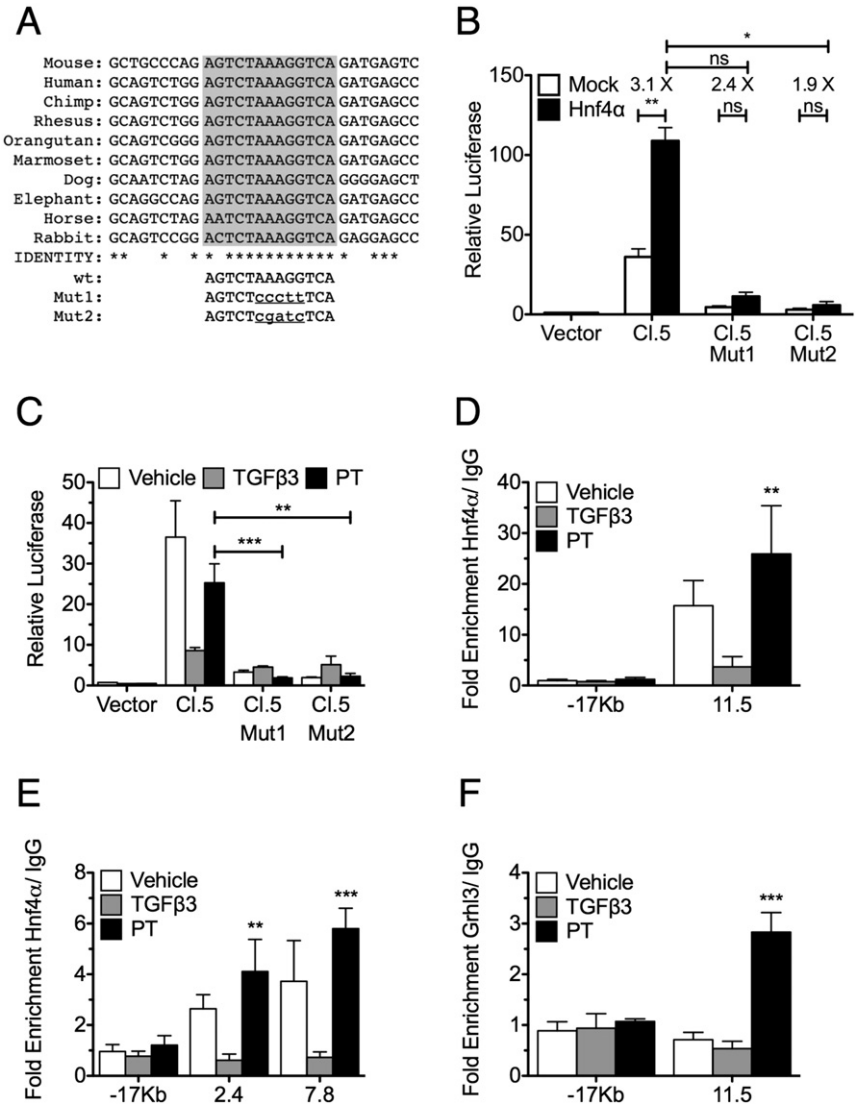

Fig. 5. Cl.5 harbors a potent and conserved Hnf4α binding site. (A) Sequence alignment of the identified Hnf4α binding site in several species. The consensus is highlighted and asterisks indicate identical bases. Introduced point mutations in the binding site are given. (B) Luciferase reporter assay in NMuMG cells transfected with reporter constructs containing the wt or mutated binding site for Hnf4α in Cl.5 with or without Hnf4α expression plasmids. (C) Mutations in the Hnf4α binding site significantly affects the basal activity of Cl.5 and abolishes changes in activity in response to TGFβ3. Relative luciferase activity is calculated as fold induction relative to empty vector control pGL4.23 (B) or to vehicle (C). (D–F) ChIP experiments with anti-Hnf4α (D and E) or anti-Grhl3 (F) antibodies in NMuMG cells in response to TGFβ3 (Supplementary Table S5). Hnf4α is enriched at the 11.5 enhancer (D) as well as at the Grhl3 bound enhancers at 2.4 and 7.8 (E) and Grhl3 is bound to the 11.5 enhancer (F).

The remaining Hnf4α-dependent increase in activity of the mutant reporter constructs was not significant.

Using ChIP experiments we detected significant enrichment of Hnf4α on this novel enhancer in NMuMG cells in a TGFβ-dependent fashion (Fig. 5D). In contrast to Grhl3 enrichment on the Cl.3 intronic sites, Hnf4α recruitment to the chromatin was detectable already in untreated NMuMG cells under steady state conditions, suggesting a distinct function for this enhancer. Hnf4α disappeared from Cl.5 during TGFβ3-mediated EMT and *Cdh1* downregulation and it was enriched 1.5 fold 72 h after TGFβ3 withdrawal (Fig. 5D). To our surprise, we also detected significant enrichment of Hnf4α at both Grhl3 enhancers at sites 2.4 and 7.8a + b (Fig. 5E, Supplementary Fig. S4C). A similar enrichment of Grhl3 was identified at the Hnf4α binding site at 11.5 (Fig. 5F). This indicated a possible interaction between these two enhancers in agreement with the 3C results that showed looping of Cl.5 to the region of Cl.3.

3.5. Grhl3 and Hnf4α functionally interact at the intronic enhancers

The Hnf4α enrichment at the Grhl3 enhancers and the interaction observed with 3C experiments intrigued us to look for a potential genetic or physical interaction between the two factors and between Cl.3 and Cl.5. First we performed luciferase reporter assays using Cl.3 and studied the effect of overexpressing Hnf4α on Grhl3 function. We observed a significant increase in reporter activity when Hnf4α was introduced together with Grhl3 compared to the Grhl3-specific activation alone (Fig. 6A). While Hnf4α alone did not increase the reporter activity, the observed effect of Hnf4α over Grhl3 was specific, since no increase was visible when we introduced Cebpα as control (data not shown).

Next, we used siRNAs to downregulate Hnf4α expression (Supplementary Fig. S4A), and found that the absence of Hnf4α resulted in a modest but significant decrease in the Grhl3-dependent activation of the reporter, compared to the control (Fig. 6B). We wanted to determine if a similar effect could be observed on *Cdh1* chromatin. For this purpose, we first treated NMuMG cells with TGFβ3 for 72 h and at the time of TGFβ3 withdrawal, cells were transfected with siRNAs targeting *Hnf4a*. After 72 h post-treatment we performed ChIP using Grhl3 specific antibodies and measured the Grhl3 enrichment at the previously identified Grhl3 sites and at the promoter. As anticipated, loss of Hnf4α expression dramatically reduced the Grhl3 enrichment at the *Cdh1* intronic enhancers as well as at the promoter (TSS), although the latter does not contain consensus sites for Hnf4α and Grhl3 (Fig. 6C, Supplementary Fig. S4E). This was accompanied by a modest but significant decrease in the levels of E-cad mRNA as measured by qPCR (Fig. 6D). We hypothesized that the Hnf4α-dependent recruitment of Grhl3 to intronic enhancers may be due to a physical interaction between the two proteins and validated this by co-immunoprecipitation experiments. We transfected HA-Hnf4α into NMuMG cells during MET after TGFβ3 removal. Anti-Grhl3 immunoprecipitation revealed a complex of Hnf4α and Grhl3 as indicated by the co-immunoprecipitation of HA-Hnf4α (Supplementary Fig. S4B). Physical interaction was also confirmed in HEK-293 cells ectopically expressing tagged versions of the two proteins. To abolish DNA-mediated indirect coupling of the two transcription factors we digested DNA by benzonase treatment of cell lysates. Immunoprecipitation of either HA-Hnf4α or flag-Grhl3 resulted in specific co-precipitation of the corresponding binding partner (Fig. 6E).

p300 is a crucial co-factor of Hnf4α-mediated gene regulation and found in the PolII-transcription initiation complex at many promoters [32,33]. p300 was detected on both enhancers in control cells and at the TSS with highest enrichment at the Grhl3 enhancer at 7.8a + b. Importantly, these levels were diminished upon depletion of Hnf4α (Fig. 6F), indicating that p300 recruitment was Hnf4α-dependent. Moreover, we addressed the effect of Hnf4α depletion on the architecture of *Cdh1* chromatin by looking at an active histone mark (acetylation at Lys9 of H3, H3K9Ac, Fig. 6G) and one that is associated with enhancers (monomethylation at Lys4 of H3, H3K4me1, Supplementary Fig. S4C and D). We found that both marks were enriched at both enhancers and at the TSS in comparison to a control region. Upon *Hnf4a* knockdown both marks were significantly reduced in those regions (Fig. 6G, Supplementary Fig. S4D). Since both Grhl3 and Hnf4α were found also at the promoter (Figs. 3A and 6C, Supplementary Fig. S4C), we investigated whether unidentified non-canonical binding sites are present at the promoter. Neither Grhl3 nor Hnf4α were increasing *Cdh1* promoter-only reporter gene activity, indicating that both factors are only indirectly linked to the TSS (Supplementary Fig. S4E). Taken together, our results suggest that Hnf4α is crucial for stabilizing the chromatin structure in a permissive state, enabling proper binding of Grhl3 to the intronic enhancers for a robust upregulation of E-cad expression.

3.6. Grhl3 contributes to the regulation of Hnf4α expression during MET

To analyze how Grhl3 and Hnf4α are connected to orchestrate the re-expression of E-cad during MET, we studied E-cad, Grhl3 and Hnf4α mRNA levels in a time course experiment after the withdrawal of TGFβ3 (Fig. 7A). We noticed that the expression of Grhl3 was restored rapidly and was first to reach 50% of its initial level within 24 h. While Hnf4α expression was also increasing, it fell behind the levels of Grhl3 and reached 50% of the initial level 6–12 h later. E-cad expression levels increased steadily but were reaching original levels later than Grhl3 and Hnf4α after 72 h (Fig. 7A). Based on these data, we hypothesized that besides its crucial role in the re-expression of E-cad, Grhl3 may also be responsible for the upregulation of Hnf4α which in turn assists Grhl3 at *Cdh1* enhancers for appropriate re-establishment of E-cad levels.

The pattern of Hnf4α expression suggested a level of regulation by Grhl3. We analyzed the expression of Hnf4α following the downregulation of Grhl3. Our results confirmed that in response to the depletion of Grhl3, a significant decrease in Hnf4α mRNA levels was observed (Fig. 7B). *Hnf4a* is activated from two alternative promoters (P1 and P2 in Fig. 7C) resulting in at least six different splice variants [34] with the major isoforms Hnf4α1 containing Exon 1A and Hnf4α7 containing Exon 1D derived from the P1 and P2 promoters, respectively [35,36]. We found a Grhl3 recognition motif in P1 at −339 bp and a very well-conserved one in P2 at −29 kb (Fig. 7C). Anti-Grhl3 ChIP revealed that Grhl3 was significantly enriched at P2 but not at P1 in cells recovering from TGFβ3, similar to the enrichment at the *Cdh1* enhancers (Fig. 7D). This indicated that isoforms containing Exon 1D are regulated by Grhl3, which is in agreement with previous findings showing that P2 is the active promoter in NMuMG cells controlling mainly Hnf4α8 expression [37]. In summary our results suggest that in response to MET signals Grhl3 and Hnf4α are immediately upregulated, recruited together with p300 to the *Cdh1* locus and cooperatively initiate E-cad expression. In parallel Grhl3 is directly activating *Hnf4a* which accelerates the MET program.

3.7. Grhl3 and Cdh1 expression is correlated in mouse mammary cells and in human breast cancer cell lines

Finally, we aimed to evaluate whether our findings are also reflected in expression data sets in mouse mammary tissue and have implications during tumor progression in human breast cancer samples. We first investigated whether *Grhl1–3*, *Hnf4a* and *Cdh1* gene expression values were correlated across normal mouse mammary tissue subpopulations (i.e., stroma, basal, luminal, epithelial and total) using data from GSE40877. We found that *Grhl1*, *Grhl2*, *Grhl3* and *Cdh1* were all highly and positively correlated with each other (Fig. S5A; Supplementary Table S8). Next, we analyzed data sets of established human breast cancer cell lines from the CCLE (Cancer Cell Line Encyclopedia) A significant positive correlation among *GRHL1–3* and *CDH1* was observed (Fig. S5B; Supplementary Table S8). Microarray data sets of a fibroblast/breast cancer cell co-culture system (GSE41678) revealed that the *CDH1* expressing Cal51 breast cancer cell line highly expressed *GRHL3* and/or *HNF4A* in monoculture and in co-culture with fibroblasts (Fig. S5C). In

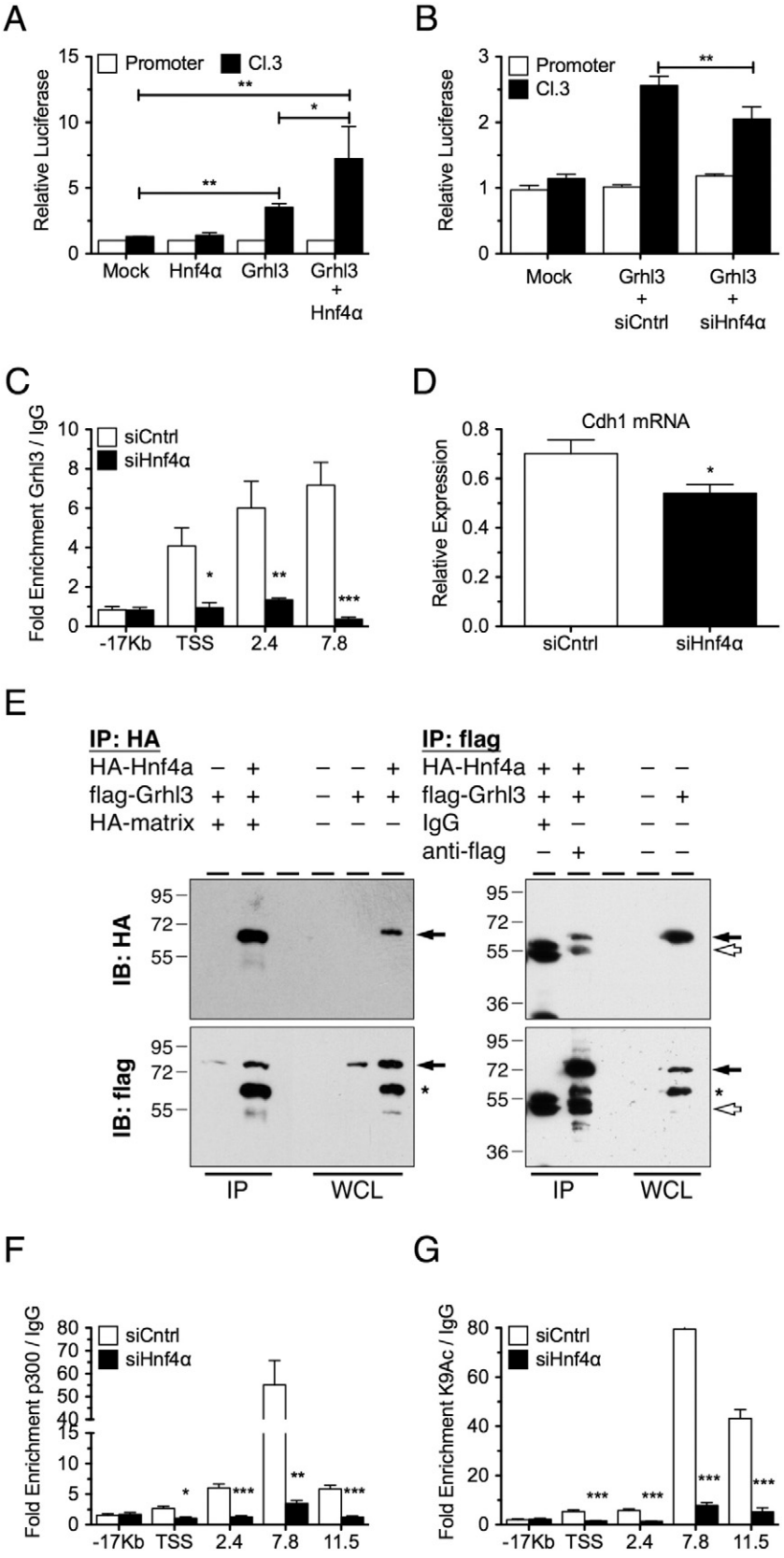

Fig. 6. Hnf4α facilitates the Grhl3 function. (A) Luciferase reporter assays in untreated NMuMG comparing the effect of Hnf4α on Grhl3 activation of Cl.3. (B) Luciferase reporter assays of untreated NMuMG cells showing the effect of Hnf4α depletion on Grhl3 activation of Cl.3. (C) ChIP of Grhl3 during MET upon the siRNA mediated downregulation of Hnf4α (Supplementary Table S5); TSS, transcriptional start site. (D) qPCR analysis of E-cad mRNA levels in NMuMG cells upon Hnf4α depletion during the recovery from TGFβ3 treatment. (E) Co-immunoprecipitation (IP) of HA-Hnf4α and flag-Grhl3 in transiently transfected HEK-293 cells showing interaction of both factors using anti-HA affinity matrix (left) and anti-Grhl3 (right) antibodies coupled to magnet beads. Immuno- and co-immunoprecipitated proteins (solid arrows) are distinguished from immunoglobulin heavy and light chains (open arrows). Bands corresponding to HA-Hnf4α derived from previous anti-HA incubation are labeled in the lower blots (asterisks). (F and G) Hnf4α is required for proper histone acetylation at the intronic enhancers. ChIP analysis of p300 (F) and H3K9Ac (G) enrichment in response to Hnf4α depletion during the recovery from TGFβ3 treatment (Supplementary Table S5).

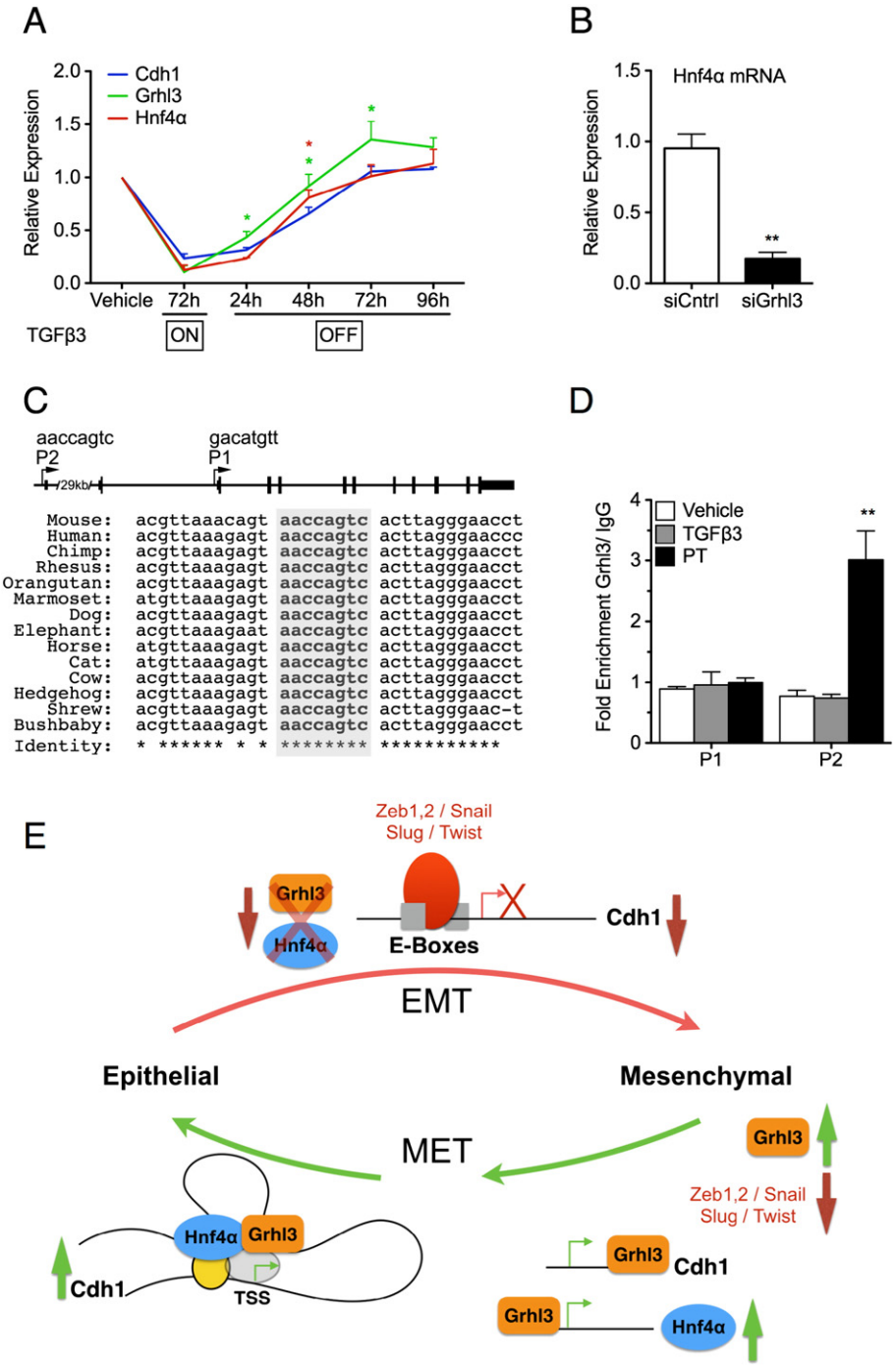

Fig. 7. Grhl3 contributes to the regulation of Hnf4α expression. (A) qPCR analysis of mRNA levels of Grhl3, Hnf4α and E-cad in response to TGFβ3 treatment and withdrawal at the indicated time-points. Green asterisks represent the p-values of comparing Grhl3 expression to that of Hnf4α and E-cad and the red ones represent the p-value of Hnf4α expression relative to E-cad. See also Supplementary Table S7. (B) qPCR analysis of Hnf4α mRNA levels upon Grhl3 knockdown by siRNAs. (C) *Hnf4a* gene structure with exons (boxes and vertical lines), showing the relative location of the two promoters (depicted with cornered arrows) and the sequence of the two Grhl3 motifs located within (top). The alignment shows the conserved Grhl3 binding site (highlighted) within the *Hnf4a* promoter P2. (D) ChIP analysis of Grhl3 enrichment on the *Hnf4a* promoters P1 and P2 in NMuMG cells in response to TGFβ3 treatment. (E) Proposed model of Grhl3 and Hnf4α mediated looping and activation of the *Cdh1* locus during MET. During EMT activators of E-cad expression are downregulated and transcription at the *Cdh1* locus is blocked by at least one of the EMT inducers Zeb1, Zeb2, Snail, Slug and Twist by binding to E-boxes at the promoter. Upon TGFβ3 withdrawal and MET induction Grhl3 is binding to sites at *Cdh1* and *Hnf4a*. Subsequent expression of Hnf4α leads to recruitment of Grhl3 and Hnf4α to intronic enhancers that induces DNA looping by interaction of the two factors and interaction at the TSS that involves PolII (gray) and p300 (yellow) (enhancer cooperativity). The assembly and stabilization of the core transcription machinery leads to induction of E-cad expression. Up- and downregulation is indicated by vertical green and red arrows, respectively.

contrast, basal like MDA-MB-231 breast cancer cells expressed all three genes at low levels regardless of the culturing methodology (Supplementary Fig. S5C). *GRHL3* expression was correlated with *GRHL1*, *GRHL2*, *CDH1* as well as with *HNF4A* (Fig. S5C; Supplementary Table S8). In line with our previous findings, these data support that *Cdh1* expression is mainly dependent on Grhl3, whereas Hnf4α might contribute to the control of the *Cdh1* locus in a tissue and context-dependent manner, especially during EMT/MET processes. In contrast to MET in NMuMG cells, based on the correlated expression in the analyzed data sets also Grhl1 and Grhl2 might be involved in *Cdh1* locus control in steady-state conditions as well.

4. Discussion

In early embryonic development proper cell adhesion mediated by E-cad is crucial for the shaping of the embryo proper and for establishing the embryonic–maternal interface of the placenta [38–41]. During these processes EMT and MET programs are initiated and marked by the switching of cadherin gene expression, suggesting that a tight and dynamic control of the *Cdh1* locus is mandatory. Proper promoter function is dependent on cis-regulatory elements located in the 47 kb intron 2.

By comparing sequences of the mammalian species mouse, man, rhesus monkey, dog and horse followed by functional analyses, we identified novel *Cdh1* enhancers at +7.8 kb and +11.5 kb downstream of the TSS. These enhancers are important for *Cdh1* gene activation during MET in NMuMG cells. Our data indicate that Grhl3 and Hnf4α are mediating this activation by binding to the enhancers at +7.8 kb and +11.5 kb, respectively. Abolishing the binding site for Grhl3 at +7.8 kb resulted in loss of expression of a *lacZ*-reporter construct in transgenic mouse embryos suggesting that this site plays a crucial role for E-cad expression in vivo as well. Moreover, Grhl3 and Hnf4α were bound to each other and cooperated during gene activation. In addition to their own binding sites they were also found to occupy each other's enhancers. Despite of the lack of binding sites within the E-cad promoter, they were also detected at the TSS, suggesting looping of DNA to complex all involved transcription factors with the basal transcription machinery (Fig. 7E).

The family of grainyhead transcription factors has been first described in Drosophila. Later three homologs of a family of six members were identified in vertebrates: grainyhead-like 1 (Grhl1), Grhl2 and Grhl3 [42–44]. In Drosophila, grainyhead (grh) is driving postembryonic neuroblast function and proliferation in part by affecting DE-cadherin expression [28,44]. In mammals, the expression is largely confined to epithelia. Although their expression only partially overlaps, overall at least one member is coexpressed with E-cad [19,42,43]. Knockout mice were described for all three, with Grhl2 showing the most severe phenotype as embryos die at E11.5 [27,29,30,45]. Each knockout show common defects in neural tube closure at specific sites, establishment of epithelial function, differentiation and barrier formation, all of which depend on proper E-cad function. Similarly, Hnf4α, an orphan member of the nuclear receptor superfamily, is also coexpressed with E-cad in several tissues such as liver, intestine and kidney [19,46]. Zygotic depletion of Hnf4α results in a pregastrulation defect due to its role in the visceral endoderm [47]. Conditional ablation of Hnf4α in the embryonic liver primordium or in the colon displays defects in epithelial specification and in organ function and homeostasis [31,48]. *Cdh1* was one prominent gene that was significantly reduced upon Hnf4α depletion in the liver, suggesting a general mechanism of *Cdh1* transcriptional control by Hnf4α in epithelia. Their role in epithelial specification, an overlapping expression with E-cad and the phenotypes of knockout mice, renders Grhl factors and Hnf4α prime candidates for the regulation of *Cdh1*.

Are Grhl factors redundant for E-cad regulation? It was recently shown by Werth and colleagues that Grhl2 acts similarly to grh in Drosophila development in regulating *Cdh1* transcription [21,28]. Loss of Grhl2 reduces E-cad expression in several epithelia including the gut tube and the otic vesicle. In mouse inner medullary collecting duct cells (mIMCD-3) and embryonic lungs Grhl2 binds to the same enhancer at 7.8a + b that was occupied by Grhl3 in our experiments [21,28]. Interestingly, Grhl2 did not show an activating effect in our system (Fig. 2B, Supplementary Fig. S3D). Only Grhl3 was able to upregulate the luciferase reporter and Grhl3 depletion during MET abolished the re-expression of E-cad and the re-establishment of an epithelial state (Fig. 4, Supplementary Fig. S3D). Although Grhl3 is highly expressed in the skin and mammary bud epithelium [19,42,43], a comprehensive expression analysis in mammary gland epithelium is lacking. Microarray re-analysis revealed that *Grhl1–3* are correlated with *Cdh1*, and thus are enriched in epithelial/luminal cell populations in mouse mammary tissues. We here used NMuMG cells mainly as a simple model for reversible EMT and MET to study *Cdh1* control using TGFβ as one of many triggers to induce EMT. Hence, although we obtained strong support of a cooperativity mechanism of the two transcription factors in *Cdh1* regulation, generalization is limited. However, also in epithelia during embryonic development and independent of MET the Grhl3-binding site is crucial for proper *Cdh1* gene activation. Since this site would potentially be bound by other factors, we cannot completely rule out that in mouse embryos our transgenic reporter construct is regulated also by other Grhl family members, conferring reduced *Cdh1* reporter gene activity independent of Grhl3. However, expression analysis showed that Grhl3 and Grhl1 rather than Grhl2 are expressed at high levels in the stomach epithelium [29,42]. Additionally, we did not observe *Cdh1* transactivation by Grhl2 in NMuMG and other cell lines, indicating tissue-specificity of individual Grhl factors. It will be interesting to analyze whether Grhl2 in mIMCD-3 cells is interacting in a similar fashion with Hnf4α or different regulators of *Cdh1* as does Grhl3.

A general regulation of *Cdh1* by Grhl3 and Hnf4α independent of MET is very likely. First, loss of Hnf4α in the liver or the intestinal epithelium of the colon substantially reduced E-cad expression due to direct interaction of Hnf4α with a site upstream of the TSS [31,48]. Second, the siRNA mediated removal of Grhl3 resulted in 30% decrease in E-cad mRNA levels also in a variety of epithelial cell lines including untreated NMuMG, CMT, CSG, P19, ESCs and keratinocytes (Supplementary Fig. S3C). Lastly, reporter gene activity in the endoderm of transgenic embryos was depending on the Grhl3 recognition motif. This indicates that Grhl3 is regulating E-cad levels not only during MET events.

Previous data suggested that the complexity of *Cdh1* gene regulation is governed by interplay of multiple cis-regulatory elements dispersed throughout the locus. We now provide evidence to how this mode of transcriptional control is achieved in molecular terms by the cooperation of two distal enhancers separated by 4 kb. Recently, a novel type of super-enhancers has been identified in embryonic stem cells. They represent clusters of normal enhancers defined by continuous several kb long stretches occupied by core transcription factors of pluripotency [49]. The presence of super-enhancers at *Cdh1* intron 2 cannot be excluded, but the detection of distinct DNaseI hypersensitive sites at each conserved cluster supports the presence of individual small enhancers (not shown). Long-range interaction of distant cis-regulatory elements and DNA looping is a general mechanism to control expression [50]. Binding to the pre-assembled transcriptional machinery is then activated by factors bound to the distal enhancers for efficient transcription [51]. Enhancer cooperativity was shown for other loci, e.g. the β-globin locus [52] and the MMP13 locus by LEF1 [53]. Using 3C analysis Vakoc et al. showed that GATA-1 and FOG-1 are required to establish the physical interaction of the locus control region and the β-major globin promoter [52]. Enhancer cooperativity during MET may reflect the need for immediate E-cad upregulation to efficiently enter an epithelial state. This enhancer cooperativity which utilizes Grhl3 and Hnf4α may also affect other epithelial genes. The detected correlation between the expression of *CDH1*, *GRHL3* and *HNF4A* in many cancer cell lines, breast cancer specimens as well as lung cancer (Supplementary Fig. S5 and data not shown), suggests a general mode of regulation. It is tempting to speculate that we identified a general mechanism that is also acting on *Cdh1* in circulating tumor cells once they activate MET and start colonization during metastasis formation.

In summary, we identified how Grhl3 is orchestrating E-cad expression during MET. External stimuli like TGFβ induce EMT leading to Hnf4α and Grhl3 downregulation and activation of transcription factors of the ZEB (Zeb1, Zeb2), Snail (Snail, Slug) and bHLH (Twist) families which results in shutdown of E-cad transcription. Upon reversion of the process e.g. by withdrawal of TGFβ EMT transcription factors are downregulated and Grhl3 expression is initiated. To facilitate MET Grhl3 acts on both *Hnf4a* and *Cdh1* transcriptional activation by binding

to the corresponding binding sites. To fully activate E-cad expression, Hnf4α is recruited to the binding site at 11.5 of the *Cdh1* locus. Here, it forms a complex with Grhl3, p300 and probably with other proteins to connect proximal and distal cis-regulatory elements by induction of a DNA loop. This facilitates the assembly of a functional transcription machinery complex at the promoter. Hnf4α and Grhl3 cooperatively bind to their cognate binding sites supporting each other's recruitment. Once the MET process is completed, Grhl3 is released from the *Hnf4*a promoter and also partially from the *Cdh1* promoter at site 7.8b and steady-state E-cad expression requires only Hnf4α (Fig. 7E).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagrm.2015.01.005.

Funding

This work was supported by the Max-Planck Society and the Deutsche Forschungsgemeinschaft SFB850 TP A4.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Acknowledgements

The Hnf4α expression construct was kindly provided by Mary Weiss. We thank Robert Kuhnert, Kati Hansen and Jessica Pfannstiel for excellent technical assistance. We are grateful to Drs. Rolf Kemler, Andreas Hecht, Jochen Maurer, Mehmet Ozturk and Uygar Tazebay for critically reading the manuscript and helpful discussions.

References

[1] J.P. Thiery, H. Acloque, R.Y. Huang, M.A. Nieto, Epithelial–mesenchymal transitions in development and disease, Cell 139 (2009) 871–890.
[2] T. Brabletz, To differentiate or not—routes towards metastasis, Nat. Rev. Cancer 12 (2012) 425–436.
[3] R. Kalluri, R.A. Weinberg, The basics of epithelial–mesenchymal transition, J. Clin. Invest. 119 (2009) 1420–1428.
[4] M.P. Stemmler, Cadherins in development and cancer, Mol. BioSyst. 4 (2008) 835–850.
[5] B. Li, Y.W. Zheng, Y. Sano, H. Taniguchi, Evidence for mesenchymal–epithelial transition associated with mouse hepatic stem cell differentiation, PLoS ONE 6 (2011) e17092.
[6] D. Vestweber, R. Kemler, P. Ekblom, Cell-adhesion molecule uvomorulin during kidney development, Dev. Biol. 112 (1985) 213–221.
[7] T. Brabletz, A. Jung, S. Reu, M. Porzner, F. Hlubek, L.A. Kunz-Schughart, R. Knuechel, T. Kirchner, Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10356–10361.
[8] R. Li, J. Liang, S. Ni, T. Zhou, X. Qing, H. Li, W. He, J. Chen, F. Li, Q. Zhuang, B. Qin, J. Xu, W. Li, J. Yang, Y. Gan, D. Qin, S. Feng, H. Song, D. Yang, B. Zhang, L. Zeng, L. Lai, M.A. Esteban, D. Pei, A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts, Cell Stem Cell 7 (2010) 51–63.
[9] T. Chen, D. Yuan, B. Wei, J. Jiang, J. Kang, K. Ling, Y. Gu, J. Li, L. Xiao, G. Pei, E-cadherin-mediated cell–cell contact is critical for induced pluripotent stem cell generation, Stem Cells 28 (2010) 1315–1325.
[10] T. Redmer, S. Diecke, T. Grigoryan, A. Quiroga-Negreira, W. Birchmeier, D. Besser, E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming, EMBO Rep. 12 (2011) 720–726.
[11] I. Bedzhov, H. Alotaibi, M.F. Basilicata, K. Ahlborn, E. Liszewska, T. Brabletz, M.P. Stemmler, Adhesion, but not a specific cadherin code, is indispensable for ES cell and induced pluripotency, Stem Cell Res. 11 (2013) 1250–1263.
[12] E. Batlle, E. Sancho, C. Franci, D. Dominguez, M. Monfar, J. Baulida, A. Garcia De Herreros, The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells, Nat. Cell Biol. 2 (2000) 84–89.
[13] J. Yang, S.A. Mani, J.L. Donaher, S. Ramaswamy, R.A. Itzykson, C. Come, P. Savagner, I. Gitelman, A. Richardson, R.A. Weinberg, Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis, Cell 117 (2004) 927–939.
[14] A. Cano, M.A. Perez-Moreno, I. Rodrigo, A. Locascio, M.J. Blanco, M.G. del Barrio, F. Portillo, M.A. Nieto, The transcription factor snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression, Nat. Cell Biol. 2 (2000) 76–83.
[15] V. Bolos, H. Peinado, M.A. Perez-Moreno, M.F. Fraga, M. Esteller, A. Cano, The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors, J. Cell Sci. 116 (2003) 499–511.
[16] J. Comijn, G. Berx, P. Vermassen, K. Verschueren, L. van Grunsven, E. Bruyneel, M. Mareel, D. Huylebroeck, F. van Roy, The two-handed E-box binding zinc finger protein Sip1 downregulates E-cadherin and induces invasion, Mol. Cell 7 (2001) 1267–1278.
[17] E.A. Carver, R. Jiang, Y. Lan, K.F. Oram, T. Gridley, The mouse snail gene encodes a key regulator of the epithelial–mesenchymal transition, Mol. Cell. Biol. 21 (2001) 8184–8188.
[18] A. Eger, K. Aigner, S. Sonderegger, B. Dampier, S. Oehler, M. Schreiber, G. Berx, A. Cano, H. Beug, R. Foisner, DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells, Oncogene 24 (2005) 2375–2385.
[19] M.P. Stemmler, A. Hecht, R. Kemler, E-cadherin intron 2 contains cis-regulatory elements essential for gene expression, Development 132 (2005) 965–976.
[20] M.P. Stemmler, A. Hecht, B. Kinzel, R. Kemler, Analysis of regulatory elements of E-cadherin with reporter gene constructs in transgenic mouse embryos, Dev. Dyn. 227 (2003) 238–245.
[21] M. Werth, K. Walentin, A. Aue, J. Schonheit, A. Wuebken, N. Pode-Shakked, L. Vilianovitch, B. Erdmann, B. Dekel, M. Bader, J. Barasch, F. Rosenbauer, F.C. Luft, K.M. Schmidt-Ott, The transcription factor grainyhead-like 2 regulates the molecular composition of the epithelial apical junctional complex, Development 137 (2010) 3835–3845.
[22] H. Hagege, P. Klous, C. Braem, E. Splinter, J. Dekker, G. Cathala, W. de Laat, T. Forne, Quantitative analysis of chromosome conformation capture assays (3C-qPCR), Nat. Protoc. 2 (2007) 1722–1733.
[23] H. Alotaibi, E. Yaman, D. Salvatore, V. Di Dato, P. Telkoparan, R. Di Lauro, U.H. Tazebay, Intronic elements in the Na+/I- symporter gene (NIS) interact with retinoic acid receptors and mediate initiation of transcription, Nucleic Acids Res. 38 (2010) 3172–3185.
[24] D. Boffelli, J. McAuliffe, D. Ovcharenko, K.D. Lewis, I. Ovcharenko, L. Pachter, E.M. Rubin, Phylogenetic shadowing of primate sequences to find functional regions of the human genome, Science 299 (2003) 1391–1394.
[25] C. Mayor, M. Brudno, J.R. Schwartz, A. Poliakov, E.M. Rubin, K.A. Frazer, L.S. Pachter, I. Dubchak, VISTA: visualizing global DNA sequence alignments of arbitrary length, Bioinformatics 16 (2000) 1046–1047.
[26] K. Quandt, K. Frech, H. Karas, E. Wingender, T. Werner, MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data, Nucleic Acids Res. 23 (1995) 4878–4884.
[27] Y. Rifat, V. Parekh, T. Wilanowski, N.R. Hislop, A. Auden, S.B. Ting, J.M. Cunningham, S.M. Jane, Regional neural tube closure defined by the Grainy head-like transcription factors, Dev. Biol. 345 (2010) 237–245.
[28] M.S. Almeida, S.J. Bray, Regulation of post-embryonic neuroblasts by Drosophila Grainyhead, Mech. Dev. 122 (2005) 1282–1293.
[29] S.B. Ting, T. Wilanowski, A. Auden, M. Hall, A.K. Voss, T. Thomas, V. Parekh, J.M. Cunningham, S.M. Jane, Inositol- and folate-resistant neural tube defects in mice lacking the epithelial-specific factor Grhl-3, Nat. Med. 9 (2003) 1513–1519.
[30] T. Wilanowski, J. Caddy, S.B. Ting, N.R. Hislop, L. Cerruti, A. Auden, L.L. Zhao, S. Asquith, S. Ellis, R. Sinclair, J.M. Cunningham, S.M. Jane, Perturbed desmosomal cadherin expression in grainy head-like 1-null mice, EMBO J. 27 (2008) 886–897.
[31] M.A. Battle, G. Konopka, F. Parviz, A.L. Gaggl, C. Yang, F.M. Sladek, S.A. Duncan, Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 8419–8424.
[32] J.C. Wang, J.M. Stafford, D.K. Granner, SRC-1 and GRIP1 coactivate transcription with hepatocyte nuclear factor 4, J. Biol. Chem. 273 (1998) 30847–30850.
[33] M.J. Barrero, S. Malik, Two functional modes of a nuclear receptor-recruited arginine methyltransferase in transcriptional activation, Mol. Cell 24 (2006) 233–243.
[34] D. Yusuf, S.L. Butland, M.I. Swanson, E. Bolotin, A. Ticoll, W.A. Cheung, X.Y. Zhang, C.T. Dickman, D.L. Fulton, J.S. Lim, J.M. Schnabl, O.H. Ramos, M. Vasseur-Cognet, C.N. de Leeuw, E.M. Simpson, G.U. Ryffel, E.W. Lam, R. Kist, M.S. Wilson, R. Marco-Ferreres, J.J. Brosens, L.L. Beccari, P. Bovolenta, B.A. Benayoun, L.J. Monteiro, H.D. Schwenen, L. Grontved, E. Wederell, S. Mandrup, R.A. Veitia, H. Chakravarthy, P.A. Hoodless, M.M. Mancarelli, B.E. Torbett, A.H. Banham, S.P. Reddy, R.L. Cullum, M. Liedtke, M.P. Tschan, M. Vaz, A. Rizzino, M. Zannini, S. Frietze, P.J. Farnham, A. Eijkelenboom, P.J. Brown, D. Laperriere, D. Leprince, T. de Cristofaro, K.L. Prince, M. Putker, L. del Peso, G. Camenisch, R.H. Wenger, M. Mikula, M. Rozendaal, S. Mader, J. Ostrowski, S.J. Rhodes, C. Van Rechem, G. Boulay, S.W. Olechnowicz, M.B. Breslin, M.S. Lan, K.K. Nanan, M. Wegner, J. Hou, R.D. Mullen, S.C. Colvin, P.J. Noy, C.F. Webb, M.E. Witek, S. Ferrell, J.M. Daniel, J. Park, S.A. Waldman, D.J. Peet, M. Taggart, P.S. Jayaraman, J.J. Karrich, B. Blom, F. Vesuna, H. O'Geen, Y. Sun, R.M. Gronostajski, M.W. Woodcroft, M.R. Hough, E. Chen, G.N. Europe-Finner, M. Karolczak-Bayatti, J. Bailey, O. Hankinson, V. Raman, D.P. LeBrun, S. Biswal, C.J. Harvey, J.P. DeBruyne, J.B. Hogenesch, R.F. Hevner, C. Heligon, X.M. Luo, M.C. Blank, K.J. Millen, D.S. Sharlin, D. Forrest, K. Dahlman-Wright, C. Zhao, Y. Mishima, S. Sinha, R. Chakrabarti, E. Portales-Casamar, F.M. Sladek, P.H. Bradley, W.W. Wasserman, The transcription factor encyclopedia, Genome Biol. 13 (2012) R24.
[35] S.F. Boj, M. Parrizas, M.A. Maestro, J. Ferrer, A transcription factor regulatory circuit in differentiated pancreatic cells, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 14481–14486.
[36] H. Nakhei, A. Lingott, I. Lemm, G.U. Ryffel, An alternative splice variant of the tissue specific transcription factor HNF4alpha predominates in undifferentiated murine cell types, Nucleic Acids Res. 26 (1998) 497–504.
[37] F. Ishikawa, K. Nose, M. Shibanuma, Downregulation of hepatocyte nuclear factor-4alpha and its role in regulation of gene expression by TGF-beta in mammary epithelial cells, Exp. Cell Res. 314 (2008) 2131–2140.
[38] I. Bedzhov, E. Liszewska, B. Kanzler, M.P. Stemmler, Igf1r signaling is indispensable for preimplantation development and is activated via a novel function of E-cadherin, PLoS Genet. 8 (2012) e1002609.

[39] L. Larue, M. Ohsugi, J. Hirchenhain, R. Kemler, E-cadherin null mutant embryos fail to form a trophectoderm epithelium, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 8263–8267.
[40] D. Riethmacher, V. Brinkmann, C. Birchmeier, A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 855–859.
[41] M.P. Stemmler, I. Bedzhov, A Cdh1HA knock-in allele rescues the Cdh1 −/− phenotype but shows essential Cdh1 function during placentation, Dev. Dyn. 239 (2010) 2330–2344.
[42] A. Auden, J. Caddy, T. Wilanowski, S.B. Ting, J.M. Cunningham, S.M. Jane, Spatial and temporal expression of the Grainyhead-like transcription factor family during murine development, Gene Expr. Patterns 6 (2006) 964–970.
[43] T. Wilanowski, A. Tuckfield, L. Cerruti, S. O'Connell, R. Saint, V. Parekh, J. Tao, J.M. Cunningham, S.M. Jane, A highly conserved novel family of mammalian developmental transcription factors related to Drosophila grainyhead, Mech. Dev. 114 (2002) 37–50.
[44] S.J. Bray, F.C. Kafatos, Developmental function of Elf-1: an essential transcription factor during embryogenesis in Drosophila, Genes Dev. 5 (1991) 1672–1683.
[45] C. Pyrgaki, A. Liu, L. Niswander, Grainyhead-like 2 regulates neural tube closure and adhesion molecule expression during neural fold fusion, Dev. Biol. 353 (2011) 38–49.
[46] S.A. Duncan, K. Manova, W.S. Chen, P. Hoodless, D.C. Weinstein, R.F. Bachvarova, J.E. Darnell Jr., Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 7598–7602.
[47] W.S. Chen, K. Manova, D.C. Weinstein, S.A. Duncan, A.S. Plump, V.R. Prezioso, R.F. Bachvarova, J.E. Darnell Jr., Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos, Genes Dev. 8 (1994) 2466–2477.
[48] W.D. Garrison, M.A. Battle, C. Yang, K.H. Kaestner, F.M. Sladek, S.A. Duncan, Hepatocyte nuclear factor 4alpha is essential for embryonic development of the mouse colon, Gastroenterology 130 (2006) 1207–1220.
[49] J. Shi, W.A. Whyte, C.J. Zepeda-Mendoza, J.P. Milazzo, C. Shen, J.S. Roe, J.L. Minder, F. Mercan, E. Wang, M.A. Eckersley-Maslin, A.E. Campbell, S. Kawaoka, S. Shareef, Z. Zhu, J. Kendall, M. Muhar, C. Haslinger, M. Yu, R.G. Roeder, M.H. Wigler, G.A. Blobel, J. Zuber, D.L. Spector, R.A. Young, C.R. Vakoc, Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation, Genes Dev. 27 (2013) 2648–2662.
[50] A. Sanyal, B.R. Lajoie, G. Jain, J. Dekker, The long-range interaction landscape of gene promoters, Nature 489 (2012) 109–113.
[51] E. Bertolino, H. Singh, POU/TBP cooperativity: a mechanism for enhancer action from a distance, Mol. Cell 10 (2002) 397–407.
[52] C.R. Vakoc, D.L. Letting, N. Gheldof, T. Sawado, M.A. Bender, M. Groudine, M.J. Weiss, J. Dekker, G.A. Blobel, Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1, Mol. Cell 17 (2005) 453–462.
[53] K. Yun, J.S. So, A. Jash, S.H. Im, Lymphoid enhancer binding factor 1 regulates transcription through gene looping, J. Immunol. 183 (2009) 5129–5137.