The class I bHLH factors E2-2A and E2-2B regulate EMT

Verónica R. Sobrado[1], Gema Moreno-Bueno[1], Eva Cubillo[1], Liam J. Holt[1,*], M. Angela Nieto[2], Francisco Portillo[1] and Amparo Cano[1,‡]

[1]Departamento de Bioquímica, Facultad de Medicina, Universidad Autónoma de Madrid (UAM), Instituto de Investigaciones Biomédicas 'Alberto Sols' (CSIC-UAM), 28029 Madrid, Spain
[2]Instituto de Neurociencias de Alicante, CSIC-UMH, 03550 Sant Joan d'Alacant, Spain
*Present address: Department of Biochemistry and Biophysics, University of California-San Francisco, San Francisco, CA 94158, USA
‡Author for correspondence (e-mail: acano@iib.uam.es)

Accepted 1 December 2008
Journal of Cell Science 122, 1014-1024 Published by The Company of Biologists 2009
doi:10.1242/jcs.028241

Summary

Functional loss of the cell-cell adhesion molecule E-cadherin is an essential event for epithelial-mesenchymal transition (EMT), a process that allows cell migration during embryonic development and tumour invasion. In most carcinomas, transcriptional repression has emerged as the main mechanism responsible for E-cadherin downregulation. Here, we report the identification of class I bHLH factor E2-2 (TCF4/ITF2) as a new EMT regulator. Both isoforms of E2-2 (E2-2A and E2-2B) induce a full EMT when overexpressed in MDCK cells but without affecting the tumorigenic properties of parental cells, in contrast to other EMT inducers, such as Snail1 or class I bHLH E47. E-cadherin repression mediated by E2-2 is indirect and independent of proximal E-boxes of the promoter. Knockdown studies indicate that E2-2 expression is dispensable for maintenance of the EMT driven by Snail1 and E47. Comparative gene-profiling analysis reveals that E2-2 factors induce similar, yet distinct, genetic programs to that induced by E47 in MDCK cells. These results, together with the embryonic expression pattern of *Tcf4* and *E2A* (which encodes E12/E47), support a distinct role for E2-2 and suggest an interesting interplay between E-cadherin repressors in the regulation of physiological and pathological EMT processes.

Supplementary material available online at
http://jcs.biologists.org/cgi/content/full/122/7/1014/DC1

Key words: bHLH E2-2, E-cadherin repression, EMT, E47, Snail, Tumour progression, Gene profiling

Introduction

Epithelial-to-mesenchymal transition (EMT) is a complex process that occurs in different biological situations involving cell migration or invasion, such as in normal embryonic development and wound healing, but also in pathologies such as cancer and organ fibrosis. During EMT, epithelial cells lose their intercellular contacts and baso-apical polarity, and acquire a fibroblastic phenotype with gain of mesenchymal markers and a motile or invasive capacity (Thiery and Sleeman, 2006). A hallmark of EMT is the functional downregulation of the cell-cell adhesion protein E-cadherin (CDH1) (Thiery, 2002; Huber et al., 2005). Diminished E-cadherin expression has been shown to correlate with increased invasiveness and metastatic potential (Birchmeier and Behrens, 1994) and to be the rate-limiting step in the transition from adenoma to carcinoma (Perl et al., 1998). E-cadherin downregulation might be transient and reversible, because the inverse process (referred to as mesenchymal-to-epithelial transition, MET) is frequently observed once metastases are established (Thiery, 2002). This concept is in accordance with the mechanisms responsible for E-cadherin downregulation in most carcinomas: promoter hypermethylation and transcriptional repression (Peinado et al., 2004a).

To increase our understanding of the molecular regulation of epithelial cell plasticity, in the last few years, our laboratory and others have focused attention on the downregulation of E-cadherin gene (*Cdh1*) expression. These studies have led to the identification of several *Cdh1* transcriptional repressors, such as the zinc-finger factors Snail1 (SNAIl) (Cano et al., 2000; Batlle et al., 2000) and Snail2 (Slug) (Hajra et al., 2002; Bolós et al., 2003), the two-handed zinc finger proteins SIP1/ZEB2 (Comijn et al., 2001) and δEF1/ZEB1 (Groteclaess and Frisch, 2000; Eger et al., 2005), and the basic helix-loop-helix (bHLH) regulators E12/E47 (Pérez-Moreno et al., 2001; Kondo et al., 2004) and Twist (Yang et al., 2004). All these factors repress E-cadherin expression, most of them through direct binding to proximal E-boxes of the promoter, and induce a full EMT when overexpressed in epithelial cells. During development, they are expressed in regions that are suffering or had undergone EMT. Importantly, their augmented expression is being detected in increasing types of tumours and invasive cell lines (for a review, see Peinado et al., 2007).

HLH transcriptional regulators are present in most eukaryotic organisms and have important roles in many essential developmental processes, regulating cell growth and differentiation of distinct cell types (Massari and Murre, 2000). All members of the family have a conserved HLH domain that mediates homo- or heterodimerisation. In addition, most HLH factors contain a stretch of basic amino acid residues, constituting a bHLH motif, through which they can bind to E-box sequences (CANNTG) in promoters of diverse genes. The HLH proteins have been classified into seven functional classes (I to VII) (Massari and Murre, 2000). In mammals, there are three class I genes: *E2A* (also known as *Tcf3*), *HEB* (*Tcf12*) and *Tcf4* (*Sef2*/*Itf2* encoding E2-2), which are also known as E-proteins and are widely expressed, but not ubiquitous. bHLH class I factors usually act as transcriptional activators, but can also function as repressors depending on the bHLH partner and/or the co-regulators they interact with (Lemercier et al., 1998; Murayama et al., 2004; Zhang et al., 2004; Goardon et al., 2006). E-proteins have two conserved transactivation regions, AD1 and AD2. The N-terminal AD1 domain seems to be required both for activation and repressor activities in different cell types (Skerjanc et al., 1996; Petropoulos and Skerjanc, 2000; Bayly et al., 2004). The

transcriptional activity of E-proteins can be negatively regulated by class V HLH factors, known as Id (inhibitor of differentiation) proteins, which lack the basic DNA-binding domain, thus blocking promoter interactions of the corresponding E-protein and Id heterodimers (Perk et al., 2005).

Class I bHLH factors have crucial roles in cell-lineage determination and lymphocyte development (Massari and Murre, 2000; Murre, 2005), although many aspects of their functionality remain unknown, which is due in part to their redundant signalling (Ik Tsen Heng and Tan, 2003). The class I *Tcf4* gene encodes two highly related isoforms, E2-2B and E2-2A, which differ in the N-terminal regions. E2-2A, the shorter isoform, contains a N-terminus devoid of the conserved AD1 domain (Henthorn et al., 1990; Corneliussen et al., 1991; Skerjanc et al., 1996). *Tcf4* is expressed in different tissues such as brain, muscle, liver, lung, testicle, trophoblasts and mammary gland, and it is implicated in T-cell and B-cell differentiation (Bergvist et al., 2000; Wikstrom et al., 2006). E2-2B has been characterized as a transcriptional repressor and a target of TCF/β-catenin (Petropoulos and Skerjanc, 2000; Kolligs et al., 2002), whereas E2-2A has been described as a transcriptional regulator with diminished or absent repressor capacity (Furumura et al., 2001; Skerjanc et al., 1996).

Here we show that E2-2 factors drive a full EMT and indirectly trigger *Cdh1* repression when overexpressed in epithelial cells. Knockdown studies indicate that, although *Tcf4* is a common downstream target of the EMT regulators Snail1, Snail2 and E47, it is dispensable for the EMT action of Snail1 and E47 factors. Nevertheless, genetic profiling and expression pattern analyses indicate that E2-2 products induce a distinct genetic program from that of E47, suggesting a non-redundant role of different bHLH factors in EMT and putting forward an interesting interplay with other *Cdh1* repressors in the regulation of EMT.

Results

E2-2 is upregulated in cells overexpressing Snail1, Snail2 or E47, and in invasive squamous carcinoma cell lines

Previous cDNA microarray studies from our laboratory revealed increased expression of the class I bHLH factor E2-2 in MDCK cells that stably express the E-cadherin repressors and EMT inducers Snail1, Snail2 or E47, relative to MDCK control cells (Moreno-Bueno et al., 2006). Those data were validated by reverse transcriptase polymerase chain reaction (RT-PCR) analysis, confirming the elevated expression of E2-2A and E2-2B mRNA transcripts in stable MDCK-Snail1, MDCK-Snail2 and MDCK-E47 cells, relative to control cells (supplementary material Fig. S1A), thus suggesting that E2-2 factors are direct or indirect targets of Snail1, Snail2 and E47. The expression of E2-2A and E2-2B was also analyzed in several keratinocyte cell lines representative of different stages of mouse skin carcinogenesis (Navarro et al., 1991; Cano et al., 2000) (supplementary material Fig. S1B). mRNA encoding E2-2B is upregulated in the invasive and E-cadherin-negative cell lines HaCa4 and CarB relative to non-invasive E-cadherin-positive MCA3D and PDV cells, whereas the E2-2A mRNA transcript was only elevated in spindle CarB cells. Interestingly, HaCa4 and CarB cells express high levels of endogenous Snail1, Snail2 and E47 factors (Cano et al., 2000; Perez-Moreno et al., 2001; Fraga et al., 2004), supporting the relationship between E2-2 and those EMT regulators inferred from the MDCK cell model. We extended the analysis of *Tcf4* expression to several human carcinoma-derived cell lines. Increased levels of transcripts encoding E2-2B were detected in SiHa and SW756 squamous cell carcinoma (SCC) cells with low or absent expression of E-cadherin, compared with well to moderately differentiated A431 (vulvar epidermoid carcinoma) and HeLa (cervical adenocarcinoma) cells (supplementary material Fig. S1C). These observations suggest that E2-2 factors, in particular E2-2B, could behave as additional *Cdh1* repressors and/or EMT inducers.

E2-2 factors induce a full EMT in epithelial cells

To test whether E2-2 factors affect E-cadherin expression and cell plasticity, the exogenous mouse E2-2B (mE2-2B) or E2-2A (mE2-2A) cDNAs were stably transfected in the prototypical MDCK epithelial cell line. Several independent stable clones were isolated and results obtained in up to three representative clones from each transfection are shown hereafter. Since the results obtained with E2-2A and E2-2B were very similar, most of the results obtained with E2-2A are shown as supplementary material. Cells transfected with mE2-2B or mE2-2A (hereafter called MDCK-E2-2B or MDCK-E2-2A, respectively) exhibited a conversion towards a spindle mesenchymal-like phenotype compared with those transfected with the control vectors (Fig. 1Aa,b; supplementary material Fig. S2A). To confirm the phenotypic change observed in these clones at the molecular level, we analyzed the expression of epithelial and mesenchymal markers. Immunofluorescence images showed drastically reduced endogenous E-cadherin protein levels and an increase or reorganization of the mesenchymal markers fibronectin and vimentin, as well as a rise in N-cadherin (Fig. 1Ac-j; supplementary material Fig. S2A). Furthermore, all clones stably expressing E2-2A and E2-2B organized F-actin in stress fibres and motile structures that were similar to lamellipodia, in contrast to the cortical F-actin organization observed in control cells (Fig. 1Ak,l; supplementary material Fig. S2A). Western blot analysis showed a decrease in β-catenin, which was similar to results in MDCK-Snail and MDCK-E47 cells (Bolós et al., 2003) (and data not shown). These experiments also confirmed the observed changes in the different epithelial and mesenchymal markers detected by immunofluorescence, with the exception of vimentin (Fig. 1B; supplementary material Fig. S2B), because no significant changes were detected at the protein level in MDCK-E2-2 clones compared with controls. Expression of the exogenous E2-2B protein in the different clones was monitored with anti-HA antibody in western blots (Fig. 1B, upper panel). In addition, immunofluorescence analysis of the tagged protein showed a nuclear localization in more than 50% of MDCK-E2-2B cells (Fig. 1C). E-cadherin downregulation was confirmed at the mRNA level and this transcript was almost undetectable in all isolated E2-2 clones (Fig. 1D; supplementary material Fig. S2C), suggesting transcriptional regulation. RT-PCR analysis revealed that expression of endogenous *Snai1* and E47 (*Tcf3*) mRNAs were not affected whereas *Snai2* transcripts were partly inhibited in MDCK-E2-2A and MDCK-E2-2B cells (Fig. 1E; supplementary material Fig. S2D). No changes in Snail1 total protein levels were either detected in MDCK-E2-2A or -E2-2B clones with respect to control MDCK-CMV cells (supplementary material Fig. S2E). Nevertheless, a potential influence of E2-2A or E2-2B factors in Snail1 stability and/or nuclear localization cannot be ruled out. Similar analyses for other EMT inducers, Snail2, E47 or ZEB1 were precluded by the lack of specific antibodies. The ability of E2-2 factors to induce EMT was confirmed in an independent cell model, the mouse keratinocyte MCA3D cell line (supplementary material Fig. S3; and data not shown). These results clearly indicate that E2-2 factors can induce

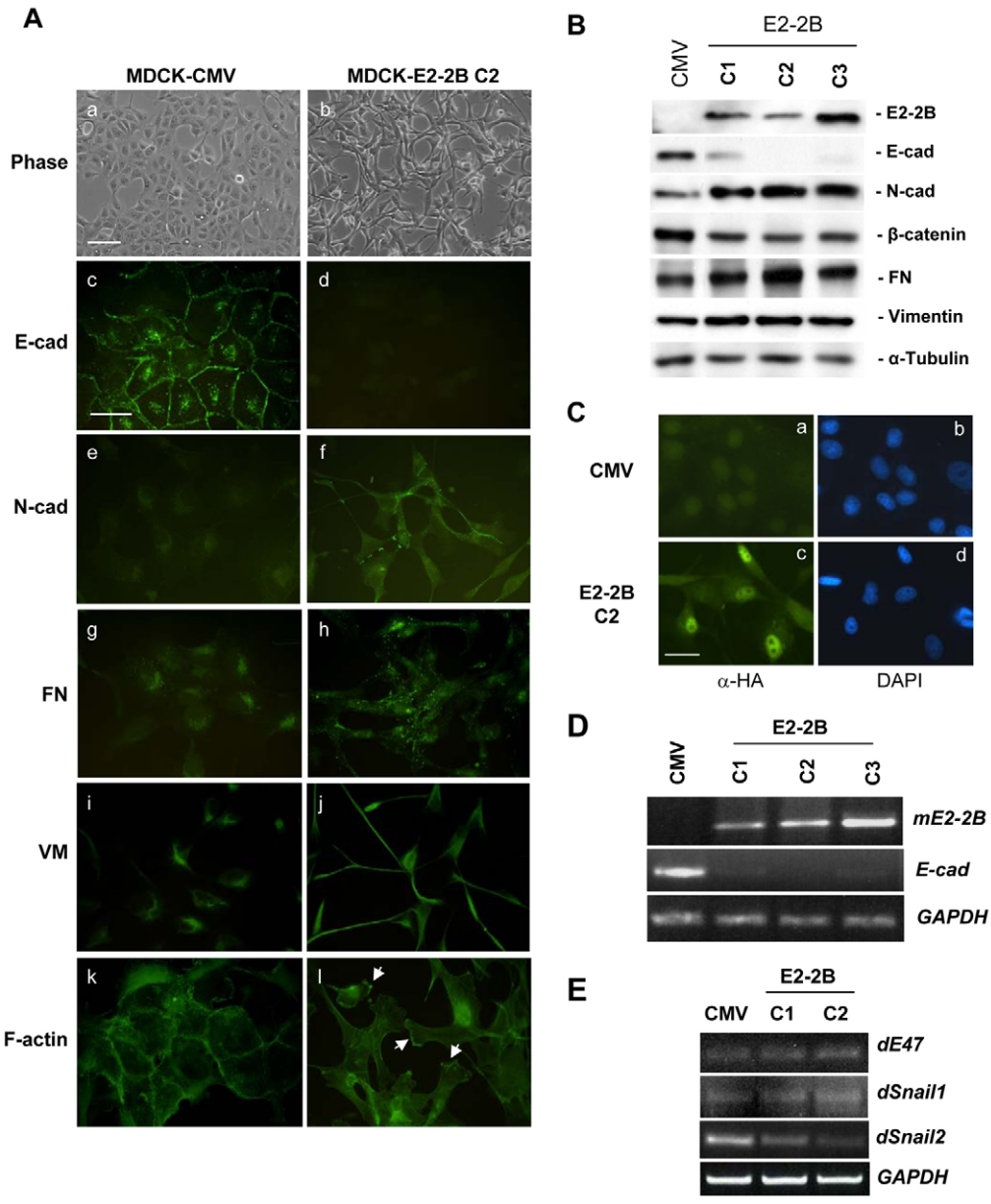

Fig. 1. E2-2B induces a full EMT when overexpressed in MDCK cells. (A) Phenotypic characterisation of MDCK-E2-2B cells. Phase-contrast images of MDCK-CMV control cells (a) and one representative clone from MDCK-E2-2B cells, C2 (b). Scale bar: 100 µm. Immunofluorescence images of EMT markers, in MDCK-CMV (c,e,g,i) and MDCK-E2-2B, C2 cells (d,f,h,j). E-cad, E-cadherin; FN, fibronectin; N-cad, N-cadherin; VM, vimentin. F-actin stain in MDCK-CMV (k) and MDCK-E2-2B (l) cells. White arrows indicate lamellipodia-like structures. Scale bar: 50 µm. (B) Western blot analysis of E2-2B-HA and the indicated epithelial and mesenchymal markers in MDCK-CMV and three independent clones from MDCK-E2-2B cells. α-Tubulin levels are shown as a loading control. (C) Immunofluorescence images showing the nuclear localization of E2-2B-HA protein. Panels a,b: control MDCK-CMV; c,d: MDCK-E2-2B, C2 cells. Scale bar: 40 µm. (D,E) Semiquantitative RT-PCR analysis to detect the expression of the exogenous E2-2B (D, upper panel) and endogenous *Cdh1* (D, middle panel) transcripts and of endogenous dog *E47*, *Snail1*and *Snail2* in the indicated clones and control cells (E). *GAPDH* levels are shown as a loading control.

a full EMT in epithelial cells with loss of the epithelial morphology and distinctive epithelial markers and gain or reorganization of mesenchymal markers.

E2-2 factors repress E-cadherin gene transcription by indirect mechanisms

The implication of E2-2 factors as E-cadherin gene (*Cdh1*) transcriptional repressors was tested on the proximal promoter of the mouse *Cdh1* gene in epithelial cells. As shown in Fig. 2A, transient transfection assays indicated that E2-2B represses *Cdh1* promoter activity by approximately 70% in epithelial MDCK cells, in a dose-dependent fashion. The same result was obtained in MCA3D and NMuMG cells (supplementary material Fig. S4A,B). Similar results were also obtained for the E2-2A isoform in MDCK cells but it failed to repress the *Cdh1* promoter activity in MCA3D and NMuMG cells (supplementary material Fig. S4C; and data not shown).

To analyze the requirement of specific E-boxes, *Cdh1* promoter constructs containing mutated versions of the promoter affecting the upstream proximal E-boxes were tested in transient transfections. As shown in Fig. 2B, the mutation of E-box1 or E-box3 does not affect the repression of the reporter by E2-2, and only a moderate de-repression was observed with mutant E-box2, indicating that proximal E-box1, E-box2 and E-box3 are essentially dispensable for E2-2 mediated *Cdh1* repression. These results suggest that E2-2 could either repress *Cdh1* promoter through binding to other regions other than E-boxes and/or in an indirect way.

Thus, to test whether E2-2 binds to the endogenous *Cdh1* promoter, chromatin immunoprecipitation (ChIP) assays were performed on MDCK-E2-2B-HA cells, using a highly specific anti-HA antibody and MDCK-Snail1-HA cells (Peinado et al., 2005) as a positive control. Surprisingly, no significant binding of E2-2B to the endogenous *Cdh1* promoter was detected in several independent clones from MDCK-E2-2B cells, in contrast to the strong Snail1

binding detected in MDCK-Snail1 cells (Fig. 2C; and data not shown). Taken together, the above results indicate that although E2-2 factors can induce an EMT process, their role in *Cdh1* repression is indirect, and probably involves the participation of another repressor(s).

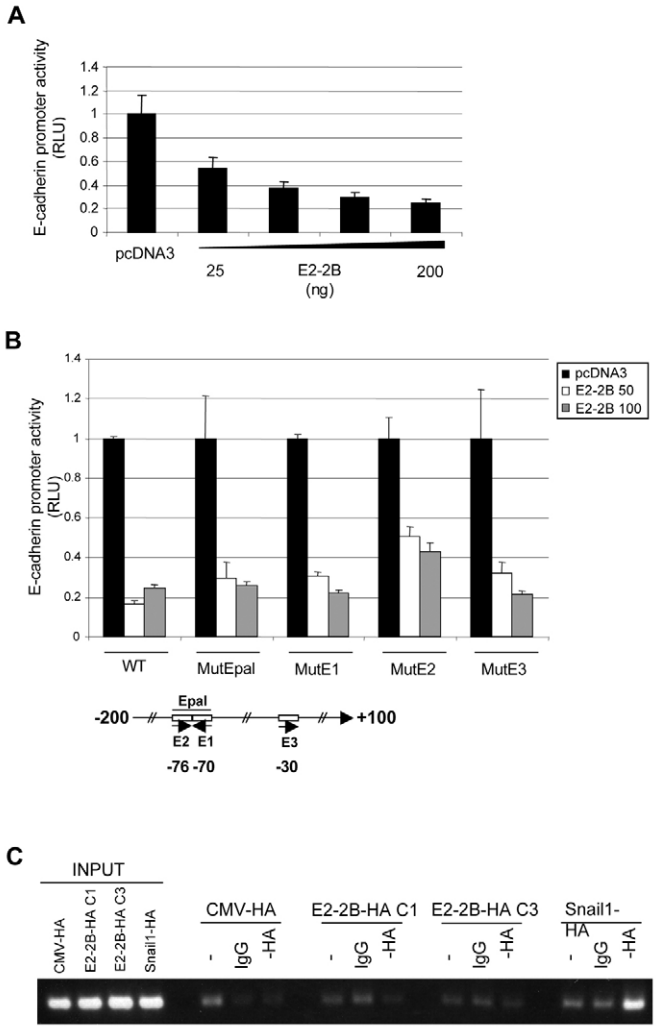

Fig. 2. E2-2B represses E-cadherin gene promoter activity in MDCK cells. (A) The activity (mean ± s.d.) of mouse proximal wild-type *Cdh1* promoter (–178/+92) fused to the luciferase gene analyzed in MDCK cells in the presence of increasing amounts (25-200 ng) of pcDNA3-E2-2B-HA or empty pcDNA3-HA control vector (200 ng). (B) The activity of the wild type (WT) *Cdh1* promoter (–178/+92) and versions containing point mutations in E-box1 (MutE1), E-box2 (MutE2), E-box3 (MutE3) or in both E-box1 and E-box2 (MutEpal) (see scheme) were analyzed in MDCK cells in the presence of pcDNA3-E2-2B-HA or empty pcDNA3-HA as indicated. CMV-β-gal was used to normalize transfection efficiency. Luciferase and β-galactosidase activities were determined 24 hours after transfection. The promoter activity is represented as relative luciferase units (RLU) and the values were normalized to those obtained in the presence of control vector. A representative experiment out of at least three experiments is shown in each panel, each performed in triplicate. Error bars represent s.d. (C) Chromatin immunoprecipitation assays performed with an anti-HA antibody in MDCK-CMV-HA, MDCK-E2-2B-HA (clones C1, C3) and MDCK-Snail1-HA cells. Amplified fragment corresponds to position –85/+130 of the dog *Cdh1* gene. Controls obtained in the absence of antibody (–) and after precipitation with an irrelevant antibody (IgG), as well as inputs, are included.

Biological relevance of E2-2 factors

The full EMT observed in MDCK clones stably expressing E2-2A and E2-2B prompted us to investigate whether they had also gained a motile or invasive behaviour. The migratory capability of MDCK-derived cells was investigated by in vitro Transwell assays on collagen type IV gels, showing that all clones overexpressing E2-2A and E2-2B displayed increased migratory ability relative to control cells (Fig. 3A). In addition, wound-healing assays on plastic dishes revealed that MDCK-E2-2A and -E2-2B cells moved faster and moved as individual cells in contrast to MDCK-CMV control cells, which migrated as an organised epithelial cell layer (data not shown). Increased motility was also observed in MCA3D-E2-2A cells in wound-healing assays (supplementary material Fig. S3C). The invasion capability of MDCK-derived cells was further

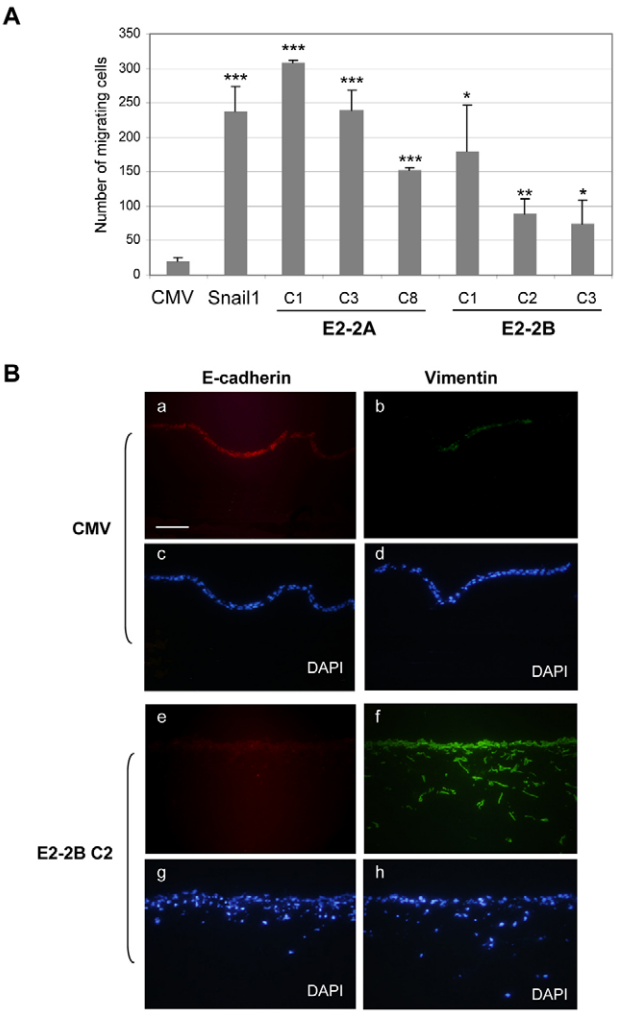

Fig. 3. E2-2B overexpression causes the acquisition of a migratory and invasive behaviour. (A) The migratory capability of MDCK-E2-2A (C1,C3,C8), MDCK-E2-2B (C1,C2,C3), MDCK-Snail1 and MDCK-CMV control cells was analyzed in Transwell assays on type IV collagen. Results represent the mean ± s.d. of three experiments performed in duplicate samples. **P*<0.05; ***P*<0.01; ****P*<0.001 in Student's *t*-test with respect to MDCK-CMV cells. (B) The infiltrating behaviour of MDCK-CMV (a-d) and MDCK-E2-2B, C2 (e-h) cells was analyzed in three-dimensional organotypic cultures on type I collagen gels. Immunofluorescence images of E-cadherin (red; a,e) and vimentin (green; b,f) from 8 µm cryostat sections from 2 week cultures. Nuclei staining with DAPI corresponding to the same fields of MDCK-CMV (c,d) and MDCK-E2-2B (g,h) cultures are also shown. Scale bar: 100 µm.

analyzed in organotypic cultures grown on three-dimensional type I collagen gels (Peinado et al., 2004b). After 2 weeks of culture, MDCK-E2-2B cells effectively invaded the collagen gel, mainly as individual cells, showed no E-cadherin expression and retained strong vimentin staining (Fig. 3Be-h). By contrast, control MDCK-CMV cells remained at the top of the gel in an organized epithelial layer (Fig. 3Ba-d), expressing E-cadherin at the basolateral membrane, as observed previously (Peinado et al., 2004b). Similar results were obtained in organotypic cultures of MDCK-E2-2A cells (supplementary material Fig. S5).

To rule out the possibility that the observed effects on migration and invasion were a consequence of increased proliferation of MDCK-E2-2 cells relative to control MDCK cells, BrdU incorporation by the different cell types was determined. All analyzed MDCK-E2-2 clones had a lower proliferation rate than control cells, both in the absence and in the presence of serum (data not shown). Thus, the motile and invasive behaviour displayed by MDCK-E2-2 cells cannot be attributed to increased proliferation.

The full EMT and invasive capacity induced by E2-2 factors in MDCK cells led us to assess the tumorigenic ability of MDCK-E2-2A and -E2-2B cells in nude mice. Unexpectedly, MDCK-E2-2A cells are nontumorigenic, whereas MDCK-E2-2B cells show very limited tumorigenic capacity (one tumour out of eight injected sites).

The motility, invasion and tumorigenic ability of MDCK-E2-2A and -E2-2B cells present two differential characteristics compared with MDCK-E47 cells, which express the closely related class I bHLH factor E47. First, MDCK-E47 cells are highly tumorigenic in nude mice and induce a potent angiogenic response (Pérez-Moreno et al., 2001; Peinado et al., 2004b). Second, MDCK-E47 cells exhibit poor infiltration ability on collagen type I organotypic cultures, and show no significant changes in their proliferation potential in two-dimensional cultures (Peinado et al., 2004b). These data indicate important differences in the biological effects of E2-2 and E47 factors and prompted us to examine the gene expression spectra induced by the expression of each factor.

Comparative genetic profile analysis between cells expressing E2-2 and E47

To examine the genetic programmes induced by each type of bHLH class I factor, we performed a comparative gene profile analysis in MDCK-E2-2 and MDCK-E47 cells using a human cDNA platform, previously validated for MDCK cells (Moreno-Bueno et al., 2006). First, we determined that no significant differences in the gene expression patterns of MDCK-E2-2A and MDCK-E2-2B clones existed (only 5 genes out of 183 were detected as differentially expressed between both cell types) (data not shown). Therefore, we compared the gene expression pattern of MDCK-E2-2A and -E2-2B (as a pool) and MDCK-E47 cells with respect to control MDCK-CMV cells. After processing the arrays, 220 clones (177 genes and 43 ESTs) were found to be modified more than twofold in expression in at least one of the two MDCK-transfected cell lines relative to the control MDCK-CMV cells (Fig. 4A; supplementary material Table S1). Around 31% (n=55) of those genes were regulated in a similar fashion (upregulated or downregulated) in both cell types (MDCK-E2-2 and MDCK-E47 cells). In addition, 12% (n=21) and 53% (n=94) of the candidate genes were modified in only one of the cell types, MDCK-E2-2 or MDCK-E47, respectively (Fig. 4B). Barely 3.6% of the clones (n=8), showed differential expression (upregulated versus downregulated, and vice versa) between MDCK-E2-2 and MDCK-E47 cells after Student's t-test analysis (FDR <0.1) (supplementary material Table S1). The functional grouping of the identified genes is shown in Fig. 4C. About 20% of the genes (35 out of 177) are related to EMT and, as expected, almost 49% of them were common to MDCK-E2-2 and MDCK-E47 cells (supplementary material Table S1). Nevertheless, a considerable fraction of EMT-related genes were specifically expressed in MDCK-E2-2 (11%, 4/35) and in MDCK-E47 cells (34%, 12/35). Importantly, analysis of the genes related to cell growth, proliferation and signalling (24% of the candidate genes) showed that only about one-third were commonly regulated in MDCK-E2-2 and MDCK-E47 cells, whereas the rest were specific to E2-2 (9.3%, 4/43) or E47 (60.4%, 26/43) cells (supplementary material Table S1). Some of the candidate genes, including those encoding SPARC and Id1, have been validated at the mRNA and/or protein level in different clones of MDCK-E2-2B and MDCK-E2-2A as well as in MDCK-E47 cells (Fig. 4D). Upregulated expression of *ZEB1* (*TCF8*) was detected in the microarray analysis in MDCK-E2-2 and MDCK-E47 cells (supplementary material Table S1) and confirmed by RT-PCR analysis (Fig. 4D), thus placing *ZEB1* as a common downstream target of both class I bHLH factors. By contrast, expression of the growth factor *PDGFC* was confirmed to be specific to MDCK-E47 cells (Fig. 4D).

Altogether, these results indicate that E2-2 and E47 factors can regulate the expression of common and specific genes that might contribute to general and particular aspects of the EMT, as previously reported for Snail1, Snail2 and E47 (Moreno-Bueno et al., 2006). Furthermore, E2-2 and E47 regulate the expression of distinct sets of genes potentially involved in the differential invasiveness and proliferation potential conferred by the expression of either type of class I bHLH factor in epithelial cells.

E2-2 is expressed in early embryonic mesoderm

The previous results suggested that E2-2 and E47 could confer different biological effects to epithelial cells during EMT. To gain further insights into the biological role of E2-2 factors, we explored the in vivo distribution of the products of the gene encoding E2-2 and its relationship with E47 expression in early mouse embryos [8.5-10 days post coitus (d.p.c.)]. In situ hybridization in whole mounted embryos and vibratome sections using a probe recognizing mRNA transcripts encoding both E2-2A and E2-2B, revealed their presence in the mesoderm, the neuroepithelium and in the branchial arches, and their absence in the heart and extraembryonic membranes, except for the allantois (Fig. 5A,C; and data not shown). Higher levels of E2-2 mRNA were found in the neural crest cells populating the branchial arches and in the limb buds (Fig. 5C). No apparent expression was detected in the embryonic epithelia regardless of their origin, either ecto or endodermal (Fig. 5E,G). These results indicate that *Tcf4* follows a similar expression pattern to that previously described for the *E2A* gene (which encodes E12/E47) (Pérez-Moreno et al., 2001) (Fig. 5B,D). However, a close inspection of the expression patterns of both genes indicated a much weaker expression of *Tcf4* transcripts in the neural tube and a more restricted pattern in the mesoderm, relative to *E2A* transcripts (compare Fig. 5E,G with 5F,H, respectively). Interestingly, the absence or extremely low levels of *Tcf4* mRNA in embryonic and extraembryonic epithelia and its expression in mesodermal tissues follows an inverse pattern with respect to *Cdh1* expression (Cano et al., 2000), supporting a repressive action in early embryos. However, their apparent absence in regions that are active in EMT (Fig. 5E), suggests a direct or indirect role for *Tcf4* products in the

maintenance of the mesenchymal phenotype in vivo, rather than in induction of EMT. Interestingly, the more restricted expression of *Tcf4* mRNA in some mesodermal populations (migratory neural crest cells or somites) compared with the expression pattern of *E2A* (Fig. 5E-H) supports a complementary, rather than redundant, role for *Tcf4* and *E2A* products in the maintenance of distinct mesenchymal cell populations in vivo.

E2-2 factors do not mediate the EMT induced by Snail or E47 factors

To further study the role of E2-2 and its relation with other EMT inducers, RNA interference experiments were performed. Two specific 19-mer oligonucleotides that target *Tcf4* mRNA transcripts were designed and cloned in pSuper-Neo-GFP vector. Both shE2-2 (shE2-2.1 and shE2-2.2) vectors inhibited mRNA encoding E2-2A and E2-2B by more than 70% after transient expression in MDCK-E2-2B cells (data not shown). Stable transfection of shE2-2.1 was then performed in MDCK-E2-2B clone C2, and cells were selected as a pool (hereafter called MDCK-E2-2B-shE2-2). In parallel, cells were transfected with an nonspecific sequence (shControl) as a control (MDCK control cells). The specificity and effectiveness of E2-2 silencing was confirmed by RT-PCR and western blot analysis of the mouse E2-2B transgene. As shown in Fig. 6A, expression of mE2-2B was strongly silenced at both mRNA and protein levels in MDCK-E2-2B-shE2-2 cells (less than 70% of the corresponding control). As expected, the high levels of endogenous *ZEB1* transcripts detected in MDCK-E2-2B cells were significantly reduced after E2-2 silencing, although they still

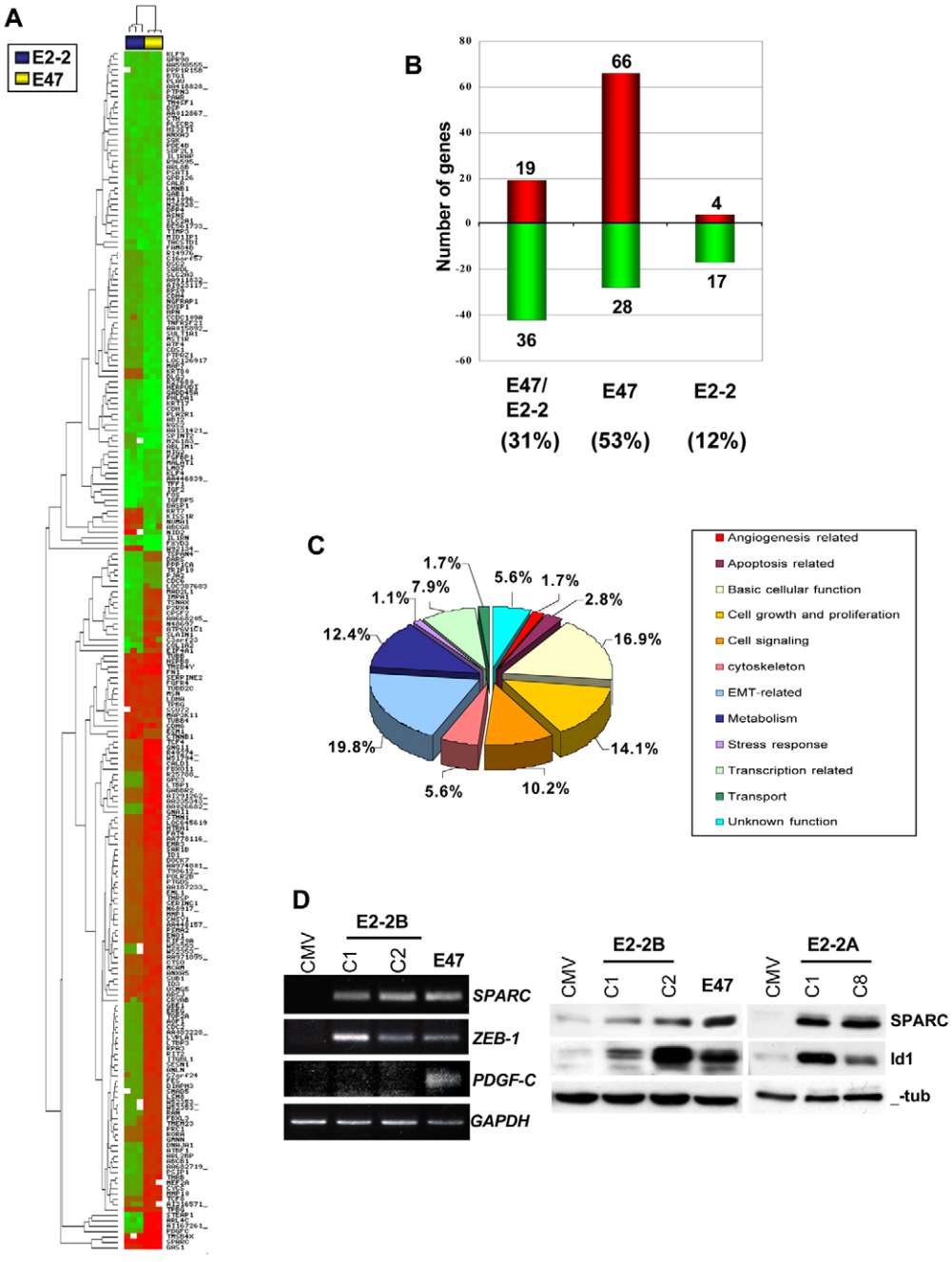

Fig. 4. Comparative gene expression analysis between MDCK-E2-2 and MDCK-E47 cells. (A) Unsupervised cluster analysis of genes modified in MDCK cells expressing E2-2 (blue) or E47 (yellow) factors, when compared with control MDCK-CMV cells. Only genes with twofold or greater variation in at least one cell type were considered and three hybridizations of each cell type were analyzed. Upregulated (red) or downregulated (green) genes are indicated. (B) Graphic representation of the common and specifically regulated genes in MDCK cells expressing E2-2 and E47. A total of 177 genes were differentially expressed more than twofold in at least one cell type relative to control cells. The number of genes found to be regulated in each case is indicated inside each bar; upregulated (red) or downregulated (green) genes are indicated. (C) Functional grouping of the modified expressed genes. The specific functions are listed on the right using a colour code and numbers indicate the percentage representation among all the candidate genes. (D) Validation of selected genes. Left, RT-PCR analysis of endogenous *SPARC*, *ZEB1* and *PDGF-C* transcripts in the indicated cell lines. Right, western blot analysis of SPARC and Id1 protein levels in the indicated cell lines. α-tubulin (right) and *GAPDH* (left) are shown as loading controls. Two independent clones from MDCK-E2-2B and MDCK-E2-2A cells were analyzed. MDCK-E47 cells were also included.

remained higher than those in MDCK-CMV control cells (Fig. 6B). However, the low endogenous levels of transcripts encoding Snail1, Snail2 and E47 were not modified in MDCK-E2-2B-shE2-2 cells compared with controls (Fig. 6B). E-cadherin analysis by RT-PCR and western blot showed that E2-2 knockdown did not restore E-cadherin expression (Fig. 6G) nor did it induce significant changes in the phenotype of MDCK-E2-2B-shE2-2 cells compared with control cells (supplementary material Fig. S6A,B), indicating that the EMT driven by E2-2 factors in MDCK cells is maintained by mechanisms that are independent of E2-2 expression.

Since *Tcf4* is a common downstream target of Snail1, Snail2 and E47 in MDCK cells (Moreno-Bueno et al., 2006), we next asked whether E2-2 is required for EMT mediated by Snail or E47 factors. Transfection of shE2-2 was performed on MDCK-Snail1 (as a model of Snail factors) and MDCK-E47 cells and sorted GFP-positive pools and independent clones were analyzed. Efficient silencing of endogenous mRNA transcripts for E2-2A and E2-2B by shE2-2 in both cell types was confirmed by RT-PCR, whereas no changes were detected in expression of the *Snai1* or *E2A* transgenes in the corresponding cells (Fig. 6C,E). Furthermore, no significant changes in endogenous levels of *Snai1*, *Snai2*, *E2A* and *ZEB1* mRNA transcripts were observed in either MDCK-Snail1 or MDCK-E47 cells following E2-2 interference (Fig. 6C-F), except for a marked decrease in *Snai2* transcripts in MDCK-E47 cells (Fig. 6D), reinforcing the specificity of E2-2 silencing. However, and similar to the findings in MDCK-E2-2B-shE2-2 cells, knockdown of E2-2 factors in MDCK-E47 or MDCK-Snail1 cells did not restore E-cadherin expression at the mRNA or protein level (Fig. 6G). Similar results were obtained in several independent clones isolated from MDCK-E47-shE2-2 cells, which also showed no changes in expression of several mesenchymal markers (N-cadherin, fibronectin, SPARC) (supplementary material Fig. S6G).

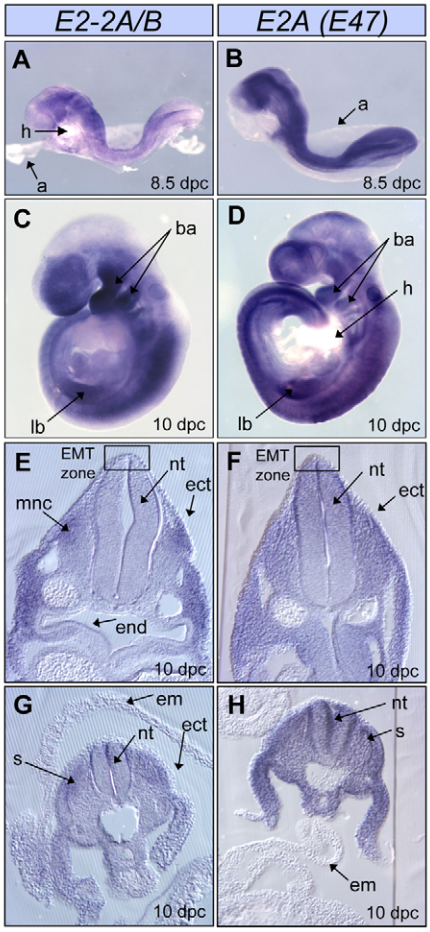

Fig. 5. *Tcf4* is expressed in the mesoderm but not in embryonic epithelia. In situ hybridization of whole-mount mouse embryos at 8.5 d.p.c. (A-B) and 10 d.p.c. (C-D), and transverse vibratome sections of 10 d.p.c. embryos taken at the level of the anterior (E-F) and posterior trunk (G-H). *Tcf4* products are detected in many different tissues throughout the embryo including the mesoderm, limb buds and neural tissues but are absent from the non-neural epithelia (E,G), the heart primordium (A) and the extra-embryonic membranes (A,G). *E2A* (*E47*) transcripts show a similar expression pattern, but with a wider distribution in the mesoderm and increased levels in the neural tube (F,H), compared with *Tcf4* transcripts (E,G). Note that the region of the neural tube undergoing EMT (EMT zone) expresses very low levels of *Tcf4* and *E2A* (*E47*) (E,F). a, amnion; ba, branchial arch; ect, ectoderm; end, endoderm; em, extraembryonic membranes; h, heart; lb, limb bud; mnc, migratory neural crest cells; nt, neural tube; s, somites.

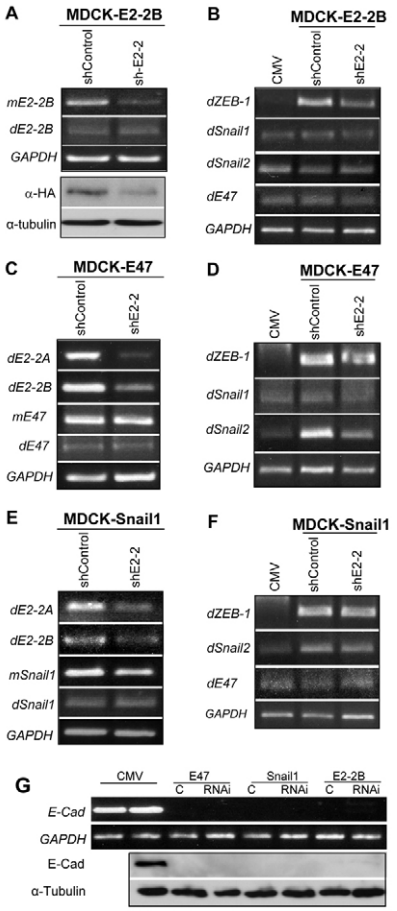

Fig. 6. Interference of *Tcf4* by shRNA. MDCK cells stably expressing E2-2B (A,B), E47 (C,D) or Snail1 (E,F) factors were transfected with shE2-2.1 or control vector and selected as pooled GFP-positive cells. Cells were analyzed by RT-PCR for expression of the indicated exogenous transgenes (A,C,D) and endogenous transcripts (A-F). Analysis of endogenous transcripts in untreated MDCK-CMV cells (B,D,F) was included as an additional control. Expression of E2-2B transgene was also confirmed by western blot analysis in MDCK-E2-2B cells (A, lower panels). (G) The expression of E-cadherin in MDCK-CMV control and indicated pooled cells after transfection with shControl (C) or shE2-2 (RNAi) was analyzed by RT-PCR (upper) and western blot (lower). *GAPDH* and α-tubulin levels were detected to control cDNA and protein loading, respectively.

Consequently, no significant changes were detected in the phenotype of MDCK-E47-shE2-2 or MDCK-Snail1-shE2-2 cells compared with controls (supplementary material Fig. S6C-F).

Taken together, the results presented here support a complementary role of E2-2 to the action of Snail and E47 factors in EMT regulation, through indirect E-cadherin repression and the induction of downstream EMT regulators, such as ZEB1.

Discussion

The study of the regulation of E-cadherin expression and EMT is relevant to the understanding of mammalian development, but more importantly, to the elucidation of the mechanisms of tumour invasion, which will hopefully facilitate the development of more-effective cancer treatments. Several EMT inducers have been identified, including the zinc-finger factors Snail1 and Snail2, the two-handed zinc-finger factors ZEB1/δEF1 and ZEB2/SIP1, and the bHLH factors Twist1 and E12/E47 (Peinado et al., 2007). Gene-profiling studies have shown that several gene clusters relevant for the EMT process lie downstream of Snail1 and Snail2, E47 or ZEB2 (De Craene et al., 2005; Vandewalle et al., 2005; Moreno-Bueno et al., 2006). Expression analysis, together with specific knockdown studies, is starting to reveal the specific participation of some of these factors in tumour progression (Yang et al., 2004; Olmeda et al., 2007a; Olmeda et al., 2007b; Olmeda et al., 2008; Aigner et al., 2007; Yuen et al., 2007; Spaderna et al., 2008). Nevertheless, the hierarchical or functional inter-relationship between different EMT regulators is still poorly understood.

The recent finding that the class I bHLH *Tcf4* gene is a downstream target of Snail1, Snail2 and E47 (Moreno-Bueno et al., 2006) led us to analyze the participation of E2-2 factors in EMT and their relationship to other EMT regulators. The results presented here demonstrate that E2-2A and E2-2B factors act as additional EMT regulators, inducing repression of E-cadherin and expression of mesenchymal markers accompanied by acquisition of motile and invasive behaviour, as exemplified in epithelial MDCK and MCA3D cell systems. In agreement with these findings, an in vivo repressor role for *Tcf4* gene products in *Cdh1* expression is supported by expression analysis carried out in early mouse embryos, showing *Tcf4* expression in those embryonic tissues where *Cdh1* expression is absent. Furthermore, the correlative data obtained in mouse and human carcinoma cells between *Tcf4* expression and an invasive or de-differentiated phenotype adds further support to the biological implication of E2-2 factors in tumour progression. In agreement with this idea, increased expression of E2-2 in a subset of breast cancer cells lines with mesenchymal traits has recently been reported (Blick et al., 2008). In this context, it is also worth mentioning that other studies have reported E2-2 upregulation in ovary adenocarcinomas with deregulated Wnt/β-catenin signalling, and revealed that the *Tcf4* gene is able to promote neoplastic transformation of RK3E cells (Zhai et al., 2002; Kolligs et al., 2002).

Detailed analysis of E2-2-mediated *Cdh1* repression indicates that *Cdh1* silencing by E2-2 factors might occur indirectly and might not involve E-boxes and their direct binding to the proximal promoter. Curiously, regarding *Cdh1* repression, the EMT inductors fall apparently into two classes: direct repressors such as Snail1, Snail2, E47, ZEB1 and ZEB2 (Batlle et al., 2000; Cano et al., 2000; Perez-Moreno et al., 2001; Comijn et al., 2001; Bolós et al., 2003; Peinado et al., 2004c; Eger et al., 2005) and indirect repressors such as Twist, FOXC2 and the homeobox protein Goosecoid (Yang and Weinberg, 2008). Thus, it seems that E2-2 belongs to this latter class. Intriguingly, maintenance of *Cdh1* repression in MDCK-E2-2 cells indeed depends on the integrity of proximal upstream E-boxes and of histone deacetylase activity (supplementary material Fig. S7), strongly suggesting the implication of downstream repressors. The analysis of the mechanistic differences between various EMT drivers deserves future investigation that will help in the elucidation of the hierarchical relationship of these factors during EMT. Nevertheless, the present results lead us to propose that E2-2 factors could induce, in certain contexts, an initial repression of *Cdh1* that will be further maintained by induction of downstream repressor/s that could also contribute to EMT maintenance. Results obtained in the knockdown studies further support this hypothesis, because no reversion of E-cadherin expression or the EMT phenotype was achieved after efficient E2-2 silencing in MDCK-E2-2 cells (Fig. 5G; supplementary material Fig. S6). This is in marked contrast to knockdown of Snail1 or E47 in MDCK-Snail1 or MDCK-E47 cells, respectively, which leads to robust re-expression of E-cadherin and a full mesenchymal-to-epithelial transition (MET) (Olmeda et al., 2007a) (E.C. and A.C., unpublished results). One of the potential candidates to maintain *Cdh1* repression and EMT in stable MDCK-E2-2 cells is ZEB1, which is a strong *Cdh1* repressor in various cell systems (Groteclaess and Frisch, 2000; Eger et al., 2005). Indeed, *ZEB1* expression is strongly upregulated in MDCK-E2-2 cells and maintained at high levels even after E2-2 silencing, whereas the endogenous levels of *Snail* and *E47* transcripts are very low in parental MDCK cells and not significantly modified by stable expression of E2-2 or following its knockdown in MDCK-E2-2 cells (Fig. 6). Unfortunately, preliminary assays directed to study the binding of ZEB1 to the *Cdh1* promoter in MDCK-E2-2 cells rendered negative results (data not shown), suggesting that an unidentified factor(s) might be responsible for maintaining *Cdh1* repression and EMT in MDCK-E2-2 cells. Further studies are clearly required to solve this important issue.

The studies performed following E2-2 silencing in MDCK-Snail1 and MDCK-E47 cells also indicated that E2-2 expression appears to be dispensable for Snail1- and E47-mediated EMT, at least in MDCK cells, suggesting a potential redundancy between these factors. However, important biological differences were detected in the behaviour of MDCK cells overexpressing E47 or E2-2 factors. Thus, although MDCK-E47 cells acquire tumorigenic properties, generating undifferentiated spindle cell tumours with high angiogenic potential (Perez-Moreno et al., 2001; Peinado et al., 2004b), MDCK-E2-2 cells showed no or very low tumorigenic capacity. However, the latter cell type exhibits a pronounced invasive potential in organotypic cultures (Fig. 3), in contrast to the poor infiltration ability of MDCK-E47 cells in this assay (Peinado et al., 2004b). These results are in agreement with recent findings that support a specific functionality of the various E-proteins in certain neural progenitors (Flora et al., 2007). In agreement with the biological differences detected between MDCK-E2-2 and MDCK-E47 cells, comparative gene expression profiling showed that E2-2 and E47 induce common and specific genetic programs when overexpressed in MDCK cells, supporting a differential role for each type of class I bHLH factor in tumour progression and invasion, outside EMT induction. For instance, constitutive expression of E47 leads to the modified expression of several genes important in the cell cycle, cell proliferation or signalling, such as *Cdc2*, *Ran*, *Smad5*, *Ltbp1/3* or *PDGFC*, which are not significantly modified (or even downregulated) in cells overexpressing E2-2 factors (supplementary material Table S1), and that can potentially explain the different tumorigenic and proliferation properties

displayed by MDCK-E2-2 and MDCK-E47 cells. In this context, it is worth mentioning that our recent comparative gene expression profile analysis of MDCK-Snail1, MDCK-Snail2 and MDCK-E47 cells also revealed that those factors share multiple direct or indirect target genes, but that can also induce specific gene expression programs (Moreno-Bueno et al., 2006), and thus confer specific functional properties to the EMT and to the tumorigenic or invasive process. Together, these results provide evidence for the existence of distinct genetic programs underlying the EMT mediated by Snail (Snail1/Snail2) and class I bHLH factors (E47/E2-2). Importantly, we show here that *Tcf4* is a, direct or indirect, downstream target of Snail1, Snail2 and E47, whereas *ZEB1* lies downstream of E2-2 and E47, as well as of Snail1, in agreement with previous reports (Guaita et al., 2002). These data support a hierarchical action among different EMT regulators, as previously suggested (Peinado et al., 2004a; Peinado et al., 2007), placing Snail1 and E47 upstream of ZEB1 and E2-2, which in turn contributes to induction of ZEB1 and perhaps other EMT regulators.

A large body of evidence now indicates that the same molecules and regulatory mechanisms are utilized for EMT processes during normal embryonic development and in pathological events such as cancer progression and organ fibrosis (Cano et al., 2000; Nieto, 2002; Peinado et al., 2004a; Thiery and Sleeman, 2006; Boutet et al., 2006). The results presented here provide additional support and add several new potential players for these EMT processes. The expression pattern of *Tcf4* and the inverse relationship with *Cdh1* expression in early embryonic development argues in favour of a role for *Tcf4* gene products in the generation and/or maintenance of the mesenchymal phenotype. However, comparison of the expression patterns of *Tcf4* and *E2A* genes in early mouse embryos (Fig. 5) (Pérez-Moreno et al., 2001) support the notion of a specific and complementary role for class I bHLH factors, E47, E2-2A and E2-2B, in the maintenance of the mesenchymal phenotype, as previously suggested in neural development (reviewed by Ik Tsen Heng and Tan, 2003). The fact that E2-2A and E2-2B factors are upregulated in MDCK cells overexpressing Snail1, Snail2 or E47 (and in highly invasive tumour cells), and also that they are expressed in migratory cells in mouse embryos, suggests that E2-2A and E2-2B might participate in intermediate-to-late stages of migration or invasion. They would thus contribute to EMT maintenance through the silencing of a specific subset of epithelial genes and/or induction of downstream EMT regulators, once Snail1, Snail2 and/or E47 are expressed and have initiated the first changes of the EMT. Although further work is required to assess the specific contribution of E2-2 factors to the development of human tumours, the present data clearly support their role as additional players in epithelial cell plasticity, with important connections to other EMT regulators.

Materials and Methods

Plasmid constructs

The complete cDNA sequence of mouse E2-2A (GenBank accession number MMU16321) containing the RSRS sequence, and E2-2B (GenBank accession number MMU16322) not containing the RSRS sequence (Skerjanc et al., 1996), were initially isolated in a yeast one-hybrid screen (Cano et al., 2000; Pérez-Moreno et al., 2001), and further subcloned from the yeast pACT2 expression vector (Clontech) into pcDNA3 or pcDNA3-HA mammalian vectors (Invitrogen), respectively, under the control of the cytomegalovirus (CMV) promoter.

Cell culture and treatments

Canine MDCK-II and derived cell lines, immortalized mouse mammary NMuMG, mouse epidermal squamous HaCa4 and spindle cell carcinoma CarB cells, human cell lines A431 (vulvar epidermoid carcinoma), HeLa (cervical adenocarcinoma), SiHa and SW756 (cervical squamous cell carcinomas), were grown in Dulbecco's modified Eagle's medium (DMEM, Life Technologies). Immortalized mouse MCA3D and transformed PDV keratinocyte cell lines were grown in a mixture of DMEM and Ham's F12 medium (1:1) and in Ham's F12 medium (Life Technologies), respectively. All media were supplemented with 10% FBS (foetal bovine serum), 10 mM glutamine (Invitrogen) and 100 μg/ml ampicillin and 32 μg/ml gentamicin (Sigma). The origin and characterization of MDCK-CMV, MDCK-Snail1, MDCK-Snail2 and MDCK-E47 cells has been previously described (Cano et al., 2000; Perez-Moreno et al., 2001; Bolós et al., 2003; Peinado et al., 2004b).

Stable transfections

Stable transfectants were obtained from parental MDCK-II and MCA3D cells after transfection with pcDNA3-E2-2B-HA, pcDNA3-E2-2A, or the corresponding control vector, using Lipofectamine (Invitrogen). Cells were grown in the presence of G418 (400 μg/ml) for 3-4 weeks to select for transfected cells. Individual clones were isolated with cloning rings. At least 10 independent clones were isolated from each transfection.

RT-PCR analysis

Total RNA from the different cell lines was extracted with Trizol reagent (Invitrogen). 2 μg RNA was used in RT-PCR experiments using specific primers for transcript detection of *Cdh1*, E2-2A/B, E2-2A, E2-2B, *Snail1*, *Snail2*, *E47*, *ZEB1*, *PDGF-C*, *SPARC* and *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase). PCR products were obtained after 30-35 cycles of amplification with an annealing temperature of 57-65°C. Sequences of the primers are indicated in supplementary material Table S2. Primers for amplification of *Cdh1*, *Snail1*, *Snail2*, *E47* and *GAPDH* have been previously described (Perez-Moreno et al., 2001; Bolós et al., 2003; Martinez-Estrada et al., 2006; Moreno-Bueno et al., 2006).

Cdh1 promoter assays

The mouse wild-type *Cdh1* promoter (–178 to +92) fused to luciferase was used to determine the activity of the *Cdh1* promoter, as described previously (Bolós et al., 2003; Peinado et al., 2004c), except that CMV-βgal was used to normalize for transfection efficiency. MDCK, NMuMG and MCA3D cells, and stable MDCK-derived clones were transiently transfected with pcDNA3-E2-2B-HA, pcDNA3-E2-2A or control vectors, using Lipofectamine (Invitrogen). The ΔEpal version, in which the Epal element has been deleted (Faraldo et al., 1997) and the mutant Epal version, in which both E-boxes (E-box1 and E-box2) of the Epal have been eliminated by point mutations, were excised from the pCAT basic vector and subcloned in the pGL2 vector (Promega) (Bolós et al., 2003). Independent mutations in E-box1, E-box2 or E-box3 were performed on the wild-type promoter construct by site-directed mutagenesis following standard procedures. In the mutant E-box1 construct, the first C of the E-box1 was changed to A (**C**AGGTG to **A**AGGTG), and in the mutant E-box2, the last G of the E-box2, to T (CACCT**G** to CACCT**T**). In the mutant E-box3 construct, the two first nucleotides were changed (**CA**CCTG to **AC**CCTG). 200 ng of each promoter construct and 10 ng CMV-βgal were used for transfection. Cell extracts were obtained 24 hours after transfection and luciferase and β-gal activities were determined using a Luciferase assay system (Promega). Where indicated, trichostatin A (TSA) (Sigma), dissolved in ethanol, was added to the culture medium at 100 nM, and maintained for 7 hours after transfection; an equivalent volume of ethanol was added to control untreated cells.

Chromatin immunoprecipitation (ChIP) assays

ChIP assays were performed in MDCK-CMV-HA, MDCK-E2-2B-HA (clones C1, C2, C3) and MDCK-Snail1-HA cell lines, using formaldehyde and 10 mM dimethyl adipimidate before sonication, as previously described (Peinado et al., 2004c). For detection of interaction between the tagged factors and the endogenous *Cdh1* promoter, an anti-HA antibody (Roche) was used. A 215 bp fragment of the dog promoter (–85/+130) was amplified with the primers 5′-CCCGCCGCAGGTGCAGC-CGCAGCCAAT-3′ (direct) and 5′-GGCGCCGCCGTACCGAGGG-3′ (reverse). PCR was carried out as follows: 35 cycles at 94°C for 40 seconds, 65°C for 40 seconds, and 72°C for 40 seconds. The amplified DNA was separated on a 2% agarose gel and visualized with SyberSafe (Invitrogen).

RNA interference

For knocking down E2-2A and E2-2B expression, two 19 nucleotide homologous sequences to both isoforms encoded by mouse and canine *Tcf4* genes were generated (ACTTTCCCTAGCTCCTTCT, shE2-2.1; GATCTATTCTCCAGATCAC, shE2-2.2) and cloned in pSuper-Neo-GFP vector (OligoEngine) following the manufacturer's instructions. As a control, an unspecific sequence from *Termotoga maritima* (Qiagen) was used. MDCK-E2-2B clone C2, MDCK-Snail1 and MDCK-E47 cells were transfected with shE2-2.1 and control constructs using Lipofectamine 2000 (Invitrogen) and sorted three times by EGFP fluorescence using a FACSVantage SE (BD) flow cytometer, to obtain populations with more than 90% EGFP-positive cells. Enriched cell pools from each transfection were used for characterization. Independent clones were isolated from MDCK-shE47-shE2-2.1 GFP-sorted cells by limited dilution.

Immunofluorescence and western blot analysis

Cells grown on coverslips were fixed in methanol (–20°C, 30 seconds) and stained for the various epithelial and mesenchymal markers as previously described (Cano et al., 2000; Perez-Moreno et al., 2001; Bolós et al., 2003). The primary and secondary antibodies used are described in supplementary material Table S3. For F-actin staining, cells were fixed in 3.7% formaldehyde-0.5% Triton X-100 (15 minutes at room temperature), followed by incubation with FITC-phalloidin for 30 minutes at 37°C and washed with PBS. DAPI (4,6-diamidino-2-phenylindole; Sigma) was used for DNA staining. Slides were mounted on Mowiol and the preparations were visualized using a Zeiss Axiophot microscope equipped with epifluorescence.

For western blot analysis, whole-cell extracts were obtained using RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS), containing protease inhibitors. 30 μg total protein from each sample were loaded on 7.5, 10 or 12% SDS polyacrylamide gel electrophoresis (PAGE) gels. Transfer, blocking and incubation with the appropriate antibodies was performed as described (Cano et al., 2000; Perez-Moreno et al., 2001; Bolós et al., 2003). The primary and secondary antibodies used are described in supplementary material Table S3. For Snail1 detection, mouse monoclonal 173EC3 antibody (1:40 dilution) was used as described (Franci et al., 2006). Blots were developed using the enhanced chemiluminescence detection system (Amersham-GE healthcare).

Transwell assays

Modified Boyden chambers, Transwells (Costar) with 8 μm pore-size filters covered with type IV rat collagen (BD) were used, as previously described (Cano et al., 2000). Briefly, 4×10^5 cells of each type were seeded onto the upper compartment. 16 hours later, cells present in the low compartment and lower part of the filters were counted.

Invasion assays

Invasion assays in three-dimensional organotypic cultures on collagen gels were performed as previously described (Peinado et al., 2004b). Briefly, 1×10^6 cells were seeded onto collagen (type I) gels in filter-culture inserts (pore size 3.0 mm, polycarbonate; BD Biosciences). The cultures were immersed in DMEM and grown for 2 weeks. Medium was completely replaced every 3 days. At the end of the culture period, the gels were immediately frozen in isopentane-cooled liquid nitrogen embedded in Tissue Tek OCT Compound (Medim). 8 μm cryostat sections were obtained for immunofluorescence analysis.

Statistical analysis

For determining statistic significance, Student's *t*-test and one-way ANOVA analyses were performed using the GraphPad Prism 4.0 software.

In situ hybridization of mouse embryos

In situ hybridization analyses of whole mounted embryos and vibratome slices were performed as described (Cano et al., 2000; Perez-Moreno et al., 2001). The mouse *Tcf4* probe corresponds to a fragment of ~1.1 kb of the E2-2B cDNA sequence (from +1610 to 2712) that would recognize both E2-2B and E2-2A mRNA transcripts. For detection of *E2A* products, a probe corresponding to the complete mouse E47 cDNA was used as previously described (Perez-Moreno et al., 2001). The slices were photographed with a Leica DMR microscope under Nomarski optics.

Differential expression profile analysis

Total RNA from the different cell lines was extracted as previously described (Moreno-Bueno et al., 2006). The results presented here correspond to the following cell lines: MDCK-E2-2, MDCK-E47 and MDCK-CMV. The cDNA array chip is the CNIO Oncochip manufactured by the CNIO Genomic Unit (http://bioinfo.cnio.es/data/oncochip) that contains 11,768 clones corresponding to 9300 different human genes. All genes have been printed in duplicate to assess reproducibility.

For analysis of E2-2- and E47-expressing cells, two extractions of RNA for each condition were performed and the corresponding RNA samples were labelled with dUTP-Cy5 and hybridized against the dUTP-Cy3-labeled MDCK-CMV control cells. An additional hybridization was performed using the dye swap label; thus a total of three hybridizations were done in E2-2- and E47-expressing cells. The fluorescence-intensity measurements of each array experiment were processed using the GenePix Pro 6.0 and MS Excel programs as recently described (Moreno-Bueno et al., 2006).

For statistical analysis, genes showing at least twofold variation in expression level respect to the MDCK-CMV control were selected. A hierarchical clustering method was applied to group the genes and samples on the basis of the similarities in expression, and the unsupervised analyses were visualized using the SOTA and TreeView software assuming euclidean distances between genes (http://bioinfo.cipf.es/). Genes with differentially significant expression between E2-2 and E47 factors were identified using the POMELO-II program (http://pomelo2.bioinfo.cnio.es/).We tested the null hypothesis of equal means among the three groups using Student's *t*-test, computing *P* values using a permutation test. To select differentially expressed genes, we adjusted for multiple testing using the False Discovery Rate (FDR) method (Hochberg and Benjamini, 1990). Here we report differential expression in genes with an FDR <0.1. All the microarray raw data tables have been deposited in the Gene Expression Omnibus under the accession numbers GSE9145 (for E2-2 analysis) and GSE9147 (for E47).

We thank Vanesa Santos for excellent technical assistance and all members of A.C.'s lab for their invaluable help and discussions. This work was supported by the Spanish Ministry of Education and Science (SAF2004-00361; SAF2007-63051) and the EV (MRTN 2004-005428) to A.C., Consolider-Ingenio 2010 CDC2007-00017 to A.C. and M.A.N., and the EU (MRTN 2004-005428) to A.C. and by the Fundación Mutua Madrileña (to G.M.B.). V.R.S. was supported by an FPU fellowship from the Spanish Ministry of Education and Science. G.M.B. is a junior investigator of the Ramón y Cajal Program 2004.

References

Aigner, K., Dampier, B., Descovich, L., Mikula, M., Sultan, A., Schreiber, M., Mikulits, W., Brabletz, T., Strand, D., Obrist, P. et al. (2007). The transcription factor ZEB1 (δEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. *Oncogene* **26**, 6979-6988.

Batlle, E., Sancho, E., Franci, C., Domínguez, D., Monfar, M., Baulida, J. and Garcia De Herreros, A. (2000). The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* **2**, 84-89.

Bayly, R., Chuen, L., Currie, R. A., Hyndman, B. D., Casselman, R., Blobel, G. A. and LeBrun, D. P. (2004). E2A-PBX1 interacts directly with the KIX domain of CBP/p300 in the induction of proliferation in primary hematopoietic cells. *J. Biol. Chem.* **279**, 55362-55371.

Bergvist, I., Eriksson, M., Saarikettu, J., Eriksson, B., Corneliussen, B., Grundstrom, T. and Holmberg, D. (2000). The basic helix-loop-helix transcription factor E2-2 is involved in T lymphocyte development. *Eur. J. Immunol.* **30**, 2857-2863.

Birchmeier, W. and Behrens, J. (1994). Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. *Biochim. Biophys. Acta* **1198**, 11-26.

Blick, T., Widodo, E., Hugo, H., Waltham, M., Lenburg, M. E., Neve, R. M. and Thompson, E. W. (2008). Epithelial mesenchymal transition traits in human breast cancer cell lines. *Clin. Exp. Metastasis* **25**, 629-642.

Bolós, V., Peinado, H., Perez-Moreno, M. A., Fraga, M. F., Esteller, M. and Cano, A. (2003). The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47. *J. Cell Sci.* **116**, 499-511.

Boutet, A., De Frutos, C. A., Maxwell, P. H., Mayol, M. J., Romero, J. and Nieto, M. A. (2006). Snail activation disrupts tissue homeostasis and induce fibrosis in the adult kidney. *EMBO J.* **25**, 5603-5613.

Cano, A., Perez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., del Barrio, M. G., Portillo, F. and Nieto, M. A. (2000). The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76-83.

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroek, D. and van Roy, F. (2001). The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol. Cell* **7**, 1267-1278.

Corneliussen, B., Thornell, A., Hallberg, B. and Grundstrom, T. (1991). Helix-Loop-Helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. *J. Virol.* **11**, 6084-6093.

De Craene, B., Gilbert, B., Stove, C., Bruyneel, E., van Roy, F. and Berx, G. (2005). The transcription factor snail induces tumor cell invasion through modulation of the epithelial cell differentiation program. *Cancer Res.* **65**, 6237-6244.

Eger, A., Aigner, K., Sonderegger, S., Dampier, B., Oehler, S., Schreiber, M., Berx, G., Cano, A., Beug, H. and Foisner, R. (2005). DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene* **24**, 2375-2385.

Faraldo, M. L. M., Rodrigo, I., Behrens, J., Birchneier, W. and Cano, A. (1997). Analysis of the E-cadherin and P-cadherin promoters in mouse keratinocyte cell lines from different stages of mouse skin carcinogenesis. *Mol. Carcinog.* **20**, 33-47.

Flora, A., Garcia, J. J., Thaller, C. and Zoghbi, H. Y. (2007). The E protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proc. Natl. Acad. Sci. USA* **104**, 15382-15387.

Fraga, M. F., Herranz, M., Espada, J., Ballestar, E., Paz, M. F., Ropero, S., Erkek, E., Bozdogan, O., Peinado, H., Niveleau, A. et al. (2004). A mouse skin carcinognenesis model reflects the aberrant DNA methylation pattern of human tumors. *Cancer Res.* **64**, 5527-5534.

Franci, C., Takkunen, M., Dave, N., Alameda, F., Gomez, S., Rodríguez, R., Escriva, M., Monserrat-Sentis, B., Baro, T., Garrido, M. et al. (2006). Expression of Snail protein in tumor-stroma interface. *Oncogene* **25**, 5134-5144.

Furumura, M., Potterf, S. B., Toyofuku, K., Matsunaga, J., Muller, J. and Hearing, V. J. (2001). Involvement of ITF2 in the transcriptional regulation of melanogenic genes. *J. Biol. Chem.* **276**, 28147-28154.

Goardon, N., Lambert, J. A., Rodriguez, P., Nissaire, P., Herblot, S., Thibault, P., Dumenil, D., Strouboulis, J., Romeo, P. H. and Hoang, T. (2006). ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. *EMBO J.* **25**, 357-366.

Grooteclaes, M. L. and Frisch, S. M. (2000). Evidence for a function of CtBP in epithelial gene regulation and anoikis. *Oncogene* **19**, 3823-3828.

Guaita, S., Puig, I., Franci, C., Garrido, M., Dominguez, D., Batlle, E., Sancho, E., Dedhar, S., De Herreros, A. G. and Baulida, J. (2002). Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression. *J. Biol. Chem.* **277**, 39209-39216.

Hajra, K. M., Chen, D. Y. and Fearon, E. R. (2002). The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res.* **62**, 1613-1618.
Henthorn, P., Kiledjian, M. and Kadesch, T. (1990). Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif. *Science* **247**, 467-470.
Hochberg, Y. and Benjamini, Y. (1990). More powerful procedures for multiple significance testing. *Stat. Med.* **9**, 811-818.
Huber, M. A., Kraut, N. and Beug, H. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr. Opin. Cell Biol.* **17**, 548-558.
Ik Tsen Heng, J. and Tan, S. S. (2003). The role of class I HLH genes in neural development - have they been overlooked? *BioEssays* **25**, 709-716.
Kolligs, F. T., Nieman, M. T., Winer, I., Hu, G., Van Mater, D., Feng, Y., Smith, I. M., Wu, R., Zhai, Y., Cho, K. R. et al. (2002). ITF-2, a downstream target of the Wnt/TCF pathway, is activated in human cancers with β-catenin defects and promotes neoplastic transformation. *Cancer Cell* **1**, 145-155.
Kondo, M., Cubillo, E., Tobiume, K., Shirakihara, T., Fukuda, N., Suzuki, H., Shimizu, K., Takehara, K., Cano, A., Saitoh, M. et al. (2004). A role for Id in the regulation of TGF-beta-induced epithelial-mesenchymal transdifferentiation. *Cell Death Differ.* **11**, 1092-1101.
Lemercier, C., To, R. Q., Carrasco, R. A. and Konieczny, S. F. (1998). The basic helix-loop-helix transcription factor Mist1 functions as a transcriptional repressor of myoD. *EMBO J.* **17**, 1412-1422.
Martínez-Estrada, O. M., Cullerés, A., Soriano, F. X., Peinado, H., Bolós, V., Martínez, F. O., Reina, M., Cano, A., Fabre, M. and Vilaró, S. (2006). The transcription factors Slug and Snail act as repressors of Claudin-1 expression in epithelial cells. *Biochem. J.* **394**, 449-457.
Massari, M. E. and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eukaryotic organisms. *Mol. Cell. Biol.* **20**, 429-440.
Moreno-Bueno, G., Cubillo, E., Sarrio, D., Peinado, H., Rodriguez-Pinilla, S. M., Villa, S., Bolós, V., Jorda, M., Fabra, A., Portillo, F. et al. (2006). Genetic profiling of epithelial cells expressing E-cadherin repressors reveals a distinct role for Snail, Slug, and E47 factors in epithelial-mesenchymal transition. *Cancer Res.* **66**, 9543-9556.
Murayama, A., Kim, A. S., Yanagisawa, J., Takeyama, K. I. and Kato, S. (2004). Transrepression by liganded nuclear receptor via bHLH activator co-regulator switching. *EMBO J.* **23**, 1598-1608.
Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. *Nat. Immunol.* **6**, 1079-1086.
Navarro, P., Gomez, M., Pizarro, A., Gamallo, C., Quintanilla, M. and Cano, A. (1991). A role for the E-cadherin cell-cell adhesion molecule during tumour progression of mouse epidermal carcinogenesis. *J. Cell Biol.* **115**, 517-533.
Nieto, M. A. (2002). The snail superfamily of zinc-finger transcription factors. *Nat. Rev. Mol. Cell. Biol.* **3**, 155-166.
Olmeda, D., Jordá, M., Peinado, H., Fabra, A. and Cano, A. (2007a). Snail silencing effectively suppresses tumour growth and invasiveness. *Oncogene* **26**, 1862-1874.
Olmeda, D., Moreno-Bueno, G., Flores, J. M., Fabra, A., Portillo, F. and Cano, A. (2007b). SNAI1 is required for tumor growth and lymph node metastasis of human breast carcinoma MDA-MB-231 cells. *Cancer Res.* **67**, 11721-11731.
Olmeda, D., Montes, A., Moreno-Bueno, G., Flores, J. M., Portillo, F. and Cano, A. (2008). Snai1 and Snai2 collaborate on tumour growth and metastasis properties of mouse skin carcinoma cell lines. *Oncogene* **27**, 4690-4701.
Peinado, H., Portillo, P. and Cano, A. (2004a). Transcriptional regulation of cadherins during development and carcinogenesis. *Int. J. Dev. Biol.* **48**, 365-375.
Peinado, H., Marin, F., Cubillo, E., Stark, H.-J., Fusenig, N., Nieto, M. A. and Cano, A. (2004b). Snail and E47 repressors of E-cadherin induce distinct invasive and angiogenic properties in vivo. *J. Cell Sci.* **117**, 2827-2839.
Peinado, H., Ballestar, E., Esteller, M. and Cano, A. (2004c). Snai1 mediates E-cadherin repression by recruitment of the Sin3A/histone deacetylase 1 (HDAC 1) / HDAC2. *Mol. Cell. Biol.* **24**, 306-319.
Peinado, H., Iglesias-de la Cruz, M. C., Olmeda, D., Csiszar, K., Fong, K. S., Vega, S., Nieto, M. A., Cano, A. and Portillo, F. (2005). A molecular role for lysyl oxidase-like 2 enzyme in snail regulation and tumor progression. *EMBO J.* **24**, 3446-3458.
Peinado, H., Olmeda, D. and Cano, A. (2007). Snail, ZEB, and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat. Rev. Cancer* **7**, 415-428.
Perez-Moreno, M. A., Locascio, A., Rodrigo, I., Dhondt, G., Portillo, F., Nieto, M. A. and Cano, A. (2001). A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions. *J. Biol. Chem.* **276**, 27424-27431.
Perk, J., Iavarone, A. and Benezra, R. (2005). Id family of helix-loop-helix proteins in cancer. *Nat. Rev. Cancer* **5**, 603-614.
Perl, A. K., Wilgenbus, P., Dahl, U., Semb, H. and Christofori, G. (1998). A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* **392**, 190-193.
Petropoulos, H. and Skerjanc, I. (2000). Analysis of the inhibition of MyoD activity by ITF-2B and full-lengh E12/E47. *J. Biol. Chem.* **275**, 25095-25101.
Skerjanc, I. S., Truong, J., Filion, P. and McBurney, M. W. (1996). A splice variant of the ITF-2 transcript encodes a transcription factor that inhibits MyoD activity. *J. Biol. Chem.* **271**, 3555-3561.
Spaderna, S., Schmalhofer, O., Wahlbuhl, M., Dimmler, A., Bauer, K., Sultan, A., Hlubek, F., Jung, A., Strand, D., Eger, A. et al. (2008). The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. *Cancer Res.* **68**, 537-544.
Thiery, J. P. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2**, 442-454.
Thiery, J. P. and Sleeman, J. P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nat. Rev. Mol. Cell. Biol.* **7**, 131-142.
Vandewalle, C., Comijn, J., De Craene, B., Vermassen, P., Bruyneel, E., Andersen, H., Tulchinsky, E., Van Roy, F. and Berx, G. (2005). SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Res.* **33**, 6566-6578.
Wikstrom, I., Forssell, J., Goncalves, M., Colucci, F. and Holmberg, D. (2006). E2-2 Regulates the expansion of Pro-B Cells and Follicular versus Marginal Zone decisions. *J. Immunol.* **177**, 6723-6729.
Yang, J. and Weinberg, R. A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev. Cell* **14**, 818-829.
Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A. and Weinberg, R. A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927-939.
Yuen, H. F., Chua, C. W., Chan, Y. P., Wong, Y. C., Wang, X. and Chan, K. W. (2007). Significance of TWIST and E-cadherin expression in the metastatic progression of prostatic cancer. *Histopathology* **50**, 648-658.
Zhai, Y., Wu, R., Schwartz, D. R., Darrah, D., Reed, H., Kolligs, F. T., Nieman, M. T., Fearon, E. R. and Cho, K. R. (2002). Role of β-catenin/T-cell factor-regulated genes in ovarian endometrioid adenocarcinomas. *Am. J. Pathol.* **160**, 1229-1238.
Zhang, J., Kalkum, M., Yamamura, S., Chait, B. T. and Roeder, R. G. (2004). E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* **305**, 1286-1289.