# E-cadherin intron 2 contains cis-regulatory elements essential for gene expression

**Marc P. Stemmler[1], Andreas Hecht[2] and Rolf Kemler[1,*]**

[1]Department of Molecular Embryology, Max-Planck Institute of Immunobiology, Stuebeweg 51, D-79108 Freiburg, Germany
[2]Andreas Hecht, Institute of Molecular Medicine and Cell Science, University of Freiburg, Stefan-Meier-Strasse 17, D-79104 Freiburg, Germany
*Author for correspondence (e-mail: kemler@immunbio.mpg.de)

*Accepted 22 December 2004*

Development 132, 965-976
Published by The Company of Biologists 2005
doi:10.1242/dev.01662

## Summary

**Cadherin-mediated cell-cell adhesion plays important roles in mouse embryonic development, and changes in cadherin expression are often linked to morphogenetic events. For proper embryonic development and organ formation, the expression of E-cadherin must be tightly regulated. Dysregulated expression during tumorigenesis confers invasiveness and metastasis. Except for the E-box motifs in the E-cadherin promoter, little is known about the existence and location of cis-regulatory elements controlling E-cadherin gene expression. We have examined putative cis-regulatory elements in the E-cadherin gene and we show a pivotal role for intron 2 in activating transcription. Upon deleting the genomic intron 2 entirely, the E-cadherin locus becomes completely inactive in embryonic stem cells and during early embryonic development. Later in development, from E11.5 onwards, the locus is activated only weakly in the absence of intron 2 sequences. We demonstrate that in differentiated epithelia, intron 2 sequences are required both to initiate transcriptional activation and additionally to maintain E-cadherin expression. Detailed analysis also revealed that expression in the yolk sac is intron 2 independent, whereas expression in the lens and the salivary glands absolutely relies on cis-regulatory sequences of intron 2. Taken together, our findings reveal a complex mechanism of gene regulation, with a vital role for the large intron 2.**

Key words: Cell adhesion, Knock-in, Transcription, Gene regulation, Gene expression, Mouse embryo, LCR, Comparative genomics

## Introduction

E-cadherin-mediated cell-cell adhesion plays an important role in cell sorting, migration and tissue remodeling during several morphogenetic events in embryogenesis and organogenesis. E-cadherin protein has been well-studied during development and in adult tissues. Its altered expression often correlates with the generation of new cell types and tissues (Butz and Larue, 1995; Hatta et al., 1987; Huber et al., 1996a; Takeichi, 1988). Already during mouse preimplantation development, E-cadherin is expressed and essential for blastocyst formation (Larue et al., 1994; Riethmacher et al., 1995; Vestweber and Kemler, 1984), but subsequently cells of the trophectoderm and parietal endoderm gradually lose E-cadherin expression (Butz and Larue, 1995; Nose and Takeichi, 1986). During gastrulation, formation of mesoderm is achieved only if E-cadherin is properly downregulated in delaminating epiblast cells at the primitive streak (Butz and Larue, 1995; Carver et al., 2001; Huber et al., 1996a). Likewise, E-cadherin expression in the ectoderm is turned off at neurulation but remains high at the ectoderm-neurectoderm borders, where it is actively involved in neural tube closure (Detrick et al., 1990; Fujimori et al., 1990; Takeichi, 1988). During skin development, the formation of hair follicles involves mesenchymal-epithelial interactions to establish follicle buds (Hardy, 1992; Hogan, 1999). In this process E-cadherin becomes downregulated and is replaced by P-cadherin (Hirai et al., 1989; Jamora et al., 2003). Conversely, E-cadherin transcription is re-initiated in cells undergoing mesenchymal-epithelial transitions during kidney organogenesis and in specific areas of the developing brain, as well as in differentiated neurons (Fannon and Colman, 1996; Shimamura et al., 1992; Shimamura and Takeichi, 1992; Vestweber et al., 1985). During these events, the expression of E-cadherin is often switched between ‘on’ and ‘off’ to determine the status of daughter cells.

Downregulation of E-cadherin is also a frequent event in tumorigenesis (Berx et al., 1998; Thiery, 2002), when the epithelial cell phenotype is lost during tumor progression. In many cases, the loss of E-cadherin, either by mutation within the coding sequence or by transcriptional downregulation, is a necessary step that promotes invasiveness (Berx et al., 1998; Perl et al., 1998; Thiery, 2002).

Although much information has been gathered about E-cadherin protein during development, organogenesis and tumor formation, little is known about the trancriptional regulation of E-cadherin, particularly how expression is activated and maintained in a developmentally and cell-type-specific manner. Several transcriptional repressors, all binding to the E-cadherin promoter region, have been identified that are able to downregulate the E-cadherin gene in specific contexts. The zinc-finger proteins Snail, Slug, δEF1/ZEB-1 and Sip-1/ZEB-

2, and the basic helix-loop-helix transcription factors Twist and E12/E47 inhibit E-cadherin expression (Batlle et al., 2000; Cano et al., 2000; Carver et al., 2001; Comijn et al., 2001; Conacci-Sorrell et al., 2003; Grooteclaes and Frisch, 2000; Peinado et al., 2004; Perez-Moreno et al., 2001; Yang et al., 2004). These regulatory factors bind to a common DNA sequence known as the E-box motif, present three times in the E-cadherin promoter. In addition, mediators of Wnt signaling, namely β-catenin and Lef-1, downregulate E-cadherin in hair follicle bud formation (Jamora et al., 2003). Lef-1 binds to a single Lef/Tcf motif upstream of the E-boxes (Huber et al., 1996b). Besides these precisely defined cis-regulatory elements at the promoter, an enhancer element in intron 1 has been identified (Behrens et al., 1991; Bussemakers et al., 1994; Hennig et al., 1995; Hennig et al., 1996; Ringwald et al., 1991; Sorkin et al., 1993). Recently, we provided evidence that the above mentioned elements are insufficient to give E-cadherin-specific expression in transgenic mice (Stemmler et al., 2003). In addition, we identified sequences in the first third of intron 2 (15 kb), that conferred some cell-type-specific gene activation (Stemmler et al., 2003). Although promising, the use of large fragments of the E-cadherin gene (between –6 and +16 kb from the transcription start) still did not recapitulate the complete endogenous expression pattern, indicating that important regulatory elements were missing in this analysis. However, this work pointed to the possibility that important regulatory sequences may be located in intron 2 of the E-cadherin gene.

Here, we have investigated the function of intron 2 sequences in proper E-cadherin gene regulation by deleting the entire intron 2 of E-cadherin by gene targeting in ES cells. We show that these sequences are essential for gene activation in early embryonic development. During late embryogenesis, intron 2 strongly enhances transcription. Additionally, we show that intron 2 is required for maintenance of E-cadherin expression after initial transcriptional activation.

## Materials and methods

### Generation of targeted E-cadherin alleles

The different targeting vectors were generated by using standard techniques (Sambrook et al., 1989). For targeting vector 1 (TV1), a genomic fragment of the mouse E-cadherin gene from –0.1 kb to +11 kb relative to the transcriptional start site was combined with a promoter fragment from –1.5 kb to –0.1 kb together with a HSV-*tk* cassette. A *betageo* cassette was inserted into the ATG codon (Stemmler et al., 2003) and a *loxP* site was inserted at the *Cla*I site at +1.2 kb, at the 5′ end of intron 2 (Fig. 1). The second targeting vector (TV2) was generated based on a genomic mouse E-cadherin fragment (BS11) containing exon 3 (Ringwald et al., 1991). A *Bst*EII site 300 bp 3′ of exon 3 was used to insert a PGK-*hyg*$^r$ cassette flanked by *FRT* sites and a single *loxP* site at the 3′ end of intron 2. The homologous recombination at the start codon was achieved by electroporation of 30 μg *Swa*I-linearized TV1 DNA into $10^7$ E14.1 ES cells (Hooper et al., 1987; Kuhn et al., 1991), which were then selected with G418 (Sigma, 250 μg/ml) and Ganciclovir (Cymeven, 2 μM). A twofold enrichment of G418-resistant clones was observed upon additional selection with Ganciclovir. Resistant clones were analyzed by Southern blotting and PCR for correct homologous recombination events at the 5′ and 3′ end of the locus. One correctly recombined clone was expanded and used for a second electroporation with 30 μg *Xho*I linearized TV2 DNA and selection of transfectants with hygromycin (Calbiochem, 200 μg/ml). An analysis similar to the first gene targeting was then carried out with resistant colonies after TV2 electroporation. Double-targeted clones were analyzed using pulse-field gel electrophoresis (PFGE) for separation of large fragments (Carle et al., 1986; Chu et al., 1986; Schwartz and Cantor, 1984) and subsequent Southern blotting to identify clones with both homologous recombination events on the same chromosome. Two independent clones were injected into C57BL/6 blastocysts, and embryos were transferred into pseudopregnant NMRI females. Chimeric males, identified by their coat color, were mated to C57BL/6 females to generate an Ecad-In2floxFRT mouse strain. Crossing of Ecad-In2floxFRT mice with ACT-*Flpe* mice (Dymecki, 1996) led to a deletion of the *hyg*$^r$ cassette (Ecad-In2flox) and, with expression of CMV-*Cre* (Schwenk et al., 1995), to the removal of intron 2 (Ecad-In2floxdel). Embryos were obtained from crosses of different strains to NMRI females or from crosses of CK14-*Cre* (Hafner et al., 2004) or CK19-*Cre* (Harada et al., 1999) males to Ecad-In2flox females. Detailed information about targeting vector sequences, PCR primers and Southern blot probes is available upon request.

### β-Galactosidase reporter gene histochemistry

Embryonic stages were determined by assuming that the appearance of a vaginal plug corresponds to embryonic day 0.5. Either whole-mount embryos, isolated organs, teratomas or ES cells were fixed in PBS/1% formaldehyde/0.2% glutaraldehyde/2 mM $MgCl_2$/5 mM EGTA/0.02% NP-40 for 5-90 minutes, washed three times with PBS/0.02% NP-40 and incubated overnight in PBS/2 mM $MgCl_2$/5 mM $K_3Fe(CN)_6$/5 mM $K_4Fe(CN)_6$/0.01% sodium desoxycholate/0.02% NP-40/1 mg/ml X-gal (Whiting et al., 1991). After post-fixation with 4% PFA, some specimens were embedded in paraffin, sectioned at 7 μm, and counterstained with Eosin or Hematoxylin/Eosin (Wilkinson and Green, 1990).

### Generation of teratomas

ES cells grown on embryonic fibroblasts were trypsinized and resuspended in PBS. Of these, $10^7$ cells in a volume of 100 μl were injected peritoneally into 129/Sv mice. After 3 weeks, teratomas were isolated and stained with X-gal for β-galactosidase activity.

### Real-time quantitative RT-PCR

RNA was isolated from embryonic halves of E7.5 embryos with an RNeasy Kit (Qiagen) and from yolk sacs with RNA-Bee reagent (ams biotechnology). RNA of one or two embryos or 2 μg total RNA was used to synthesize cDNA with oligo(dT)-primer and a Superscript II Kit (Invitrogen). Amplification of *betageo* RNA was carried out with the primer pair 5′-TTACTGCCGCCTGTTTTGAC-3′ and 5′-TAGCCGAATAGCCTCTCCAC-3′, and that of *Gapd* with the primer pair 5′-ACCACAGTCCATGCCATCACT-3′ and 5′-GTCCACCACCCTGTTGCTGTA-3′ [in both cases using FastStart DNA Master$^{PLUS}$ (Roche) in the LightCycler Instrument (Roche) according to the manufacturer's instructions]. Transcripts were normalized to *Gapd* expression. Values in arbitrary units are the mean of three separate experiments comparing Ecad-In2flox and Ecad-In2floxdel samples.

## Results

### Generation of mice lacking intron 2 of the E-cadherin gene

We performed an in silico comparative genomics approach of large sequence parts, including the E-cadherin locus for mouse, rat, human, chimp and dog (see Fig. S1 in the supplementary material). No significant evolutionary conservation was detected further upstream of the previously analyzed region (–6 kb of the transcription start site) (Stemmler et al., 2003). But interestingly, several blocks of sequence conservation over all five species were identified throughout the large intron 2. This

suggested that additional, not yet functionally analyzed, sequences in intron 2 are required for proper E-cadherin gene function.

A scheme for the deletion of the entire intron 2 of the E-cadherin gene (45 kb genomic sequence) is depicted in Fig. 1A. Two independent homologous recombination events were used to insert *loxP* sites 5′ and 3′ of intron 2. Additionally, we inserted a *betageo* reporter gene at the start codon of E-cadherin to monitor the transcriptional activity of the targeted locus (TV1, Fig. 1A). More than 80% of ES-cell clones were homologously recombined (Fig. 1B) after electroporation of TV1. A 6.2 kb wild-type fragment and a 9 kb fragment of the mutated allele were detected with probe *a* in Southern blot analysis after *Bam*HI digestion (Fig. 1B). One recombined ES-cell clone was taken for the second gene targeting. The 3′ *loxP* site was inserted by homologous recombination at exon 3 with targeting vector 2 (TV2, Fig. 1A, right side). Southern blot analysis showed homologous recombination at the 3′ end of the locus with a frequency of 10% (Fig. 1C). A *Bam*HI digest probed with probe *f* revealed a 12 kb wild-type fragment and a 7 kb fragment of the mutated allele due to the insertion of a *Bam*HI site at the *loxP* site. To identify recombination events which had occurred on the same allele, pulse-field gel electrophoresis separation and Southern blot analysis were performed. Hybridization with probes *e* and *c* (Fig. 1A,D) revealed a fragment that migrates at the predicted size corresponding to recombination in cis (clones 2, 6, 8-11, arrowhead, Fig. 1D). By contrast, in addition to the wild-type fragment of ~400 kb (arrow in Fig. 1D), a fragment of ~300 kb with probe *e* (Fig. 1D, left) and of 100 kb with probe *c* (Fig. 1D, right) appeared in cases where the homologous recombination event occurred in trans (clones 3-5, 7, open arrow, Fig. 1D). Three ES-cell clones with both homologous recombination events in cis were used to generate transgenic mice. Neither a potential fused mRNA between *betageo* and E-cadherin sequences as a result of the knock-in nor a hypomorphic fusion protein was detected in heterozygous mice (data not shown). Because of the *betageo* insertion at the ATG codon of E-cadherin, the targeted allele should result in a null phenotype. Consistent with the null having an early lethal phenotype (Larue et al., 1994; Riethmacher et al., 1995), interbreeding of mice heterozygous for the targeted allele failed to generate any viable homozygous knock-in offspring (data not shown).

## Deletion of intron 2 leads to loss of reporter gene expression in ES cells

First insights into the regulatory function of sequences in intron 2 were obtained with the targeted ES cells (Ecad-In2flox), which, after transient transfection with a *Cre* expression vector (Gu et al., 1993), removed intron 2 (Ecad-In2floxdel), as demonstrated by PCR and Southern blot (Fig. 2A,B). X-Gal staining of Ecad-In2flox ES cells revealed β-galactosidase (β-gal) activity, albeit in a heterogeneous pattern (Fig. 2C). By contrast, no β-gal staining was detectable in Ecad-In2floxdel ES cells (Fig. 2D). Teratomas were produced in isogenic mice from Ecad-In2flox and Ecad-In2floxdel ES cells and in both cases these tumors contained the well-known typical variety of different tissues and cell types. Reporter gene activity was observed throughout teratomas derived from Ecad-In2flox cells (Fig. 2E) and was particularly strong in cysts and polarized epithelia (Fig. 2G). However, in teratomas derived from Ecad-In2floxdel cells, only partial and weaker β-gal expression was observed (Fig. 2F), and this did not coincide with the locations of cysts (Fig. 2H). Importantly, epithelia of Ecad-In2floxdel teratomas did not stain for β-gal (Fig. 2H). These results provide strong evidence that intron 2 is necessary for the expression of E-cadherin in ES cells and in teratoma-derived differentiated epithelia. To study the differences in gene activity that are due to the function of intron 2, we compared the abundance of *betageo* transcripts in Ecad-In2flox versus Ecad-In2floxdel ES cells using a semi-quantitative PCR approach. Transcripts for *betageo* were detected in Ecad-In2flox samples, and these were much less abundant in Ecad-In2floxdel samples (Fig. 2I, upper panel). This result was verified by quantitative PCR, which showed a 95% reduction in gene activity after deletion of intron 2 (Fig. 2I, lower panel), thus confirming the pivotal role for intron 2 in activating E-cadherin gene expression.

## Cis-regulatory elements of intron 2 are required for consistent E-cadherin gene activity during early development

Next, we analyzed the contribution of the intron 2 sequences to E-cadherin gene expression during development by crossing of Ecad-In2flox mice to a *Cre*-deleter strain. Reporter gene activity of Ecad-In2flox and Ecad-In2floxdel mice was monitored by X-gal staining on embryos of different stages. Embryos from Ecad-ATG (see Fig. S2 in the supplementary material) and Ecad-In2flox mice exhibited comparable profiles and both reporter lines reflected the endogenous E-cadherin expression pattern. Particularly at E6.5, β-gal expression was found in the extra-embryonic ectoderm in higher amounts compared with the embryonic part (Fig. 3A). β-Gal staining was increased in the embryonic part at E7.5 (Fig. 3B), downregulated in the mesoderm at gastrulation, and maintained in ectoderm and endoderm (Fig. 3F), all in accordance with the known endogenous E-cadherin expression. Intense β-gal expression was observed in the definitive gut endoderm between E8.5 and E10.5, with increasing expression in the surface ectoderm (Fig. 3C-E,G). From E8.5 onwards, expression in the yolk sac was detected, and this increased until E10.5 (Fig. 3E). Importantly, β-gal expression was not found in embryos carrying the Ecad-In2floxdel locus prior to E9.5-10.0 (Fig. 3H-K). In particular, cells or tissues positive for β-gal expression from the Ecad-In2flox allele were all negative when intron 2 was absent, e.g. the extra-embryonic ectoderm at E6.5 (compare Fig. 3A and H), ectoderm and endoderm at E7.5 (compare Fig. 3B,F with 3I,M, respectively), or definitive gut endoderm at E8.5 to E10.5 (compare Fig. 3C,D,E,G with J,K,L,N, respectively). Generally, no β-gal expression of the Ecad-In2floxdel locus was seen in most high-level E-cadherin expression domains such as the lens. Exceptions to this rule are weak activities at the apical ectodermal ridge (AER) of the forelimb buds at E10.5 (Fig. 3L) and between the first and second branchial arches around E11.0 (data not shown). Interestingly, in extra-embryonic cells of the yolk sac, β-gal activity was found at comparable levels in Ecad-In2flox and Ecad-In2floxdel embryos at E10.5. Differences in gene activity between the two alleles were examined by semi-quantitative RT-PCR from embryonic cups of E7.5 embryos and revealed reduced mRNA levels after deletion of intron 2 (Fig. 3O, upper

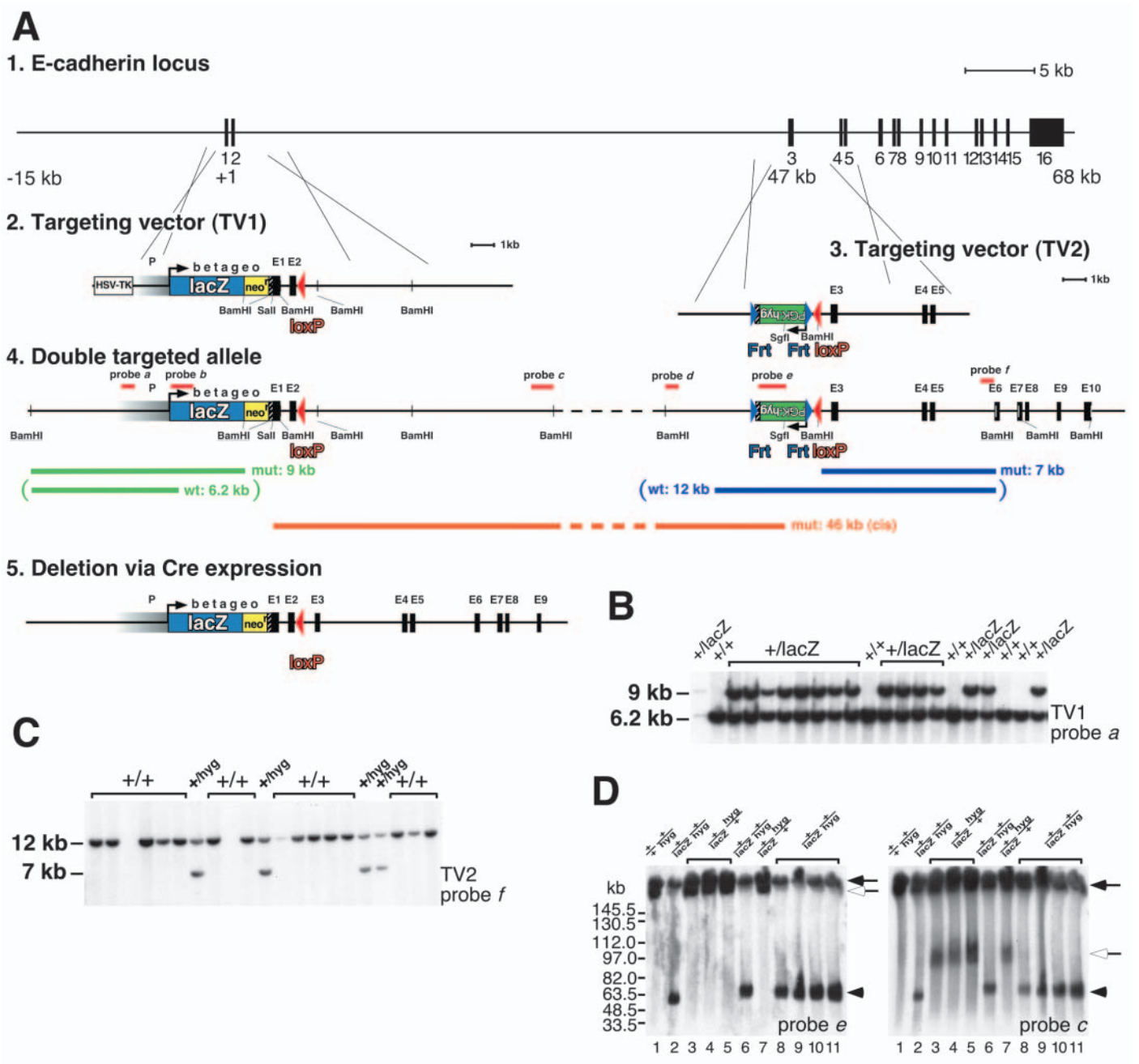


**Fig. 1.** Generation of ES cells with targeted floxed E-cadherin intron 2. (A) Schematic representation of the E-cadherin locus (drawn to scale, 1). Exons are represented by vertical black bars, and nucleotide positions are given with respect to the transcription start site (+1). The locus was targeted with vector TV1 (2) and subsequently with TV2 (3), with detailed analysis after each step, finally resulting in the double-targeted allele (4) to delete intron 2 by Cre recombinase expression (5). For additional negative selection, a herpes simplex virus thymidine kinase gene (HSV-*tk*) was integrated in TV1 and *betageo* was fused in-frame to the E-cadherin start codon. In TV2 a hygromycin resistance cassette (*hyg*$^r$) under the control of the phosphoglycerol kinase promoter (PGK) was inserted in reverse orientation 5′ of exon 3. Promoter (P), exons (E1, E2, etc.), *loxP* sites (red triangles), *FRT* sites (blue triangles), polyadenylation signals (striped boxes), transcription start sites (horizontal arrows), used restriction sites and probes (horizontal red bars) are given. The expected fragments of the Southern blot analysis for the homologous recombination of TV1 with probe *a* are indicated by green bars, and those for TV2 with probe *f* by blue bars. If both events occur at the same allele (in cis), a 46 kb fragment is expected after digestion with *Sal*I and *Sgf*I with probe *e* and with probe *c* (orange bar). (B) Southern blot analysis of *Bam*HI-digested ES-cell DNA of gene targeting with TV1 as outlined in A. A 6.2 kb fragment was observed in wild-type clones (+/+), and an additional 9.2 kb fragment in recombined clones (+/lacZ). (C) Southern blot analysis of *Bam*HI-digested ES-cell DNA of second gene targeting (TV2). Besides a 12 kb wild-type fragment, a 7 kb fragment was detected in successfully targeted clones (+/hyg). (D) Pulse-field electrophoresis separation of *Sal*I/*Sgf*I-digested ES-cell DNA of double-targeted clones analyzed by Southern blot, hybridized with probe *e* (left) or probe *c* (right). Events on the same allele are easily distinguishable by the appearance of a 46 kb fragment in both panels (arrowhead) in addition to the wild-type fragment (arrow). In clones with trans orientation, an additional fragment of >150 kb is visible with probe *e* (left, white arrow) and a different fragment of~90 kb with probe *c* (right, white arrow).

panel). Additional analysis by real-time PCR showed a 85% reduction in transcript abundancy in Ecad-In2floxdel embryos at E7.5 (Fig. 3O, lower panel). These data demonstrate that during early embryogenesis the cis-regulatory elements in intron 2 are absolutely required for gene expression from the E-cadherin locus and that the promoter alone is insufficient to drive expression.

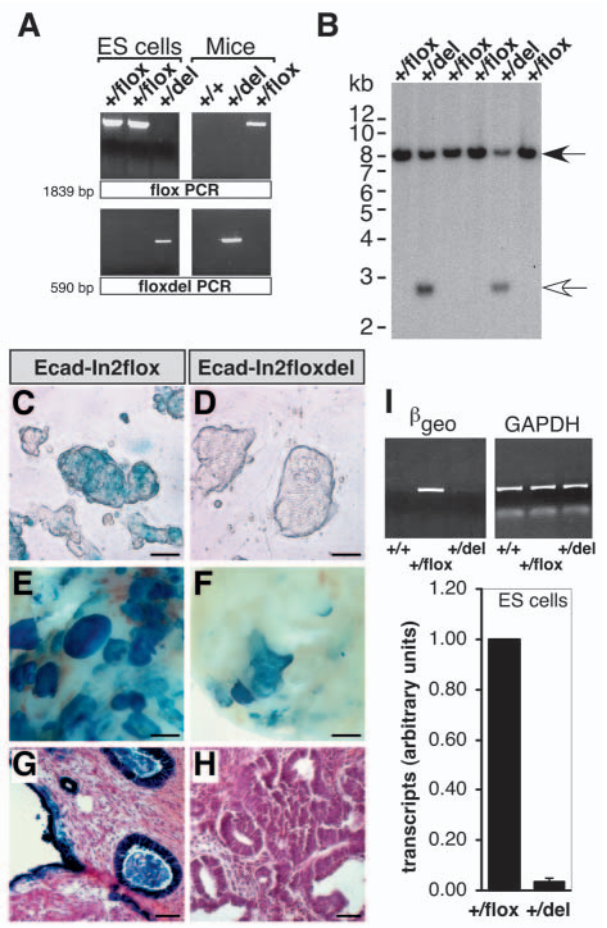


**Fig. 2.** E-cadherin-specific expression is lost in ES cells after deletion of intron 2. (A) Genotyping of Ecad-In2flox (+/flox) and Ecad-In2floxdel (+/del) ES cells after transient expression of *Cre* (left) and offspring of corresponding knock-in mouse strains (right). Primers specific for the floxed (upper panel) and floxdel allele (lower panel) were used. (B) Southern blot analysis of *Kpn*I-fragmented DNA of Ecad-In2flox ES cell clones after *Cre* expression, using a radioactive probe specific for exon 2. The restriction fragments of the wild-type and Ecad-In2flox alleles migrate at 7.7 kb (black arrow), whereas that for the Ecad-In2floxdel allele migrates at 2.7 kb (white arrow). (C,D) X-gal staining of Ecad-In2flox (C) and Ecad-In2floxdel ES cells (D) shows that expression is lost upon deletion of intron 2 sequences. (E-H) Analysis of differentiated ES cells in teratomas. Expression of β-gal is seen in teratomas from Ecad-In2flox cells (E), and this is significantly reduced in teratomas without intron 2 (F). (G,H) Sections of teratomas shown in E,F counterstained with Hematoxylin/Eosin. High-level expression in cystic epithelia in Ecad-In2flox (G) is lost in Ecad-In2floxdel teratomas. (I) Semi-quantitative (upper panel) and real-time PCR (lower panel) of both ES-cell lines with primers specific for *betageo* and *Gapd* transcripts. A reduction in gene activity is observed in the semi-quantitative and real-time PCR. Values resulting from *Gapd* real-time PCR were used for standardization. Scale bars: 50 μm in C,D,G,H; 500 μm in E,F.

## E-cadherin gene activity is significantly reduced but not lost during late embryogenesis in Ecad-In2floxdel embryos

The results of the expression analysis in early embryogenesis up to E10.5 support a pivotal role for intron 2 in establishing high-level gene activity of the E-cadherin locus. Next, we analyzed whether these cis-regulatory elements have a similar important function later in development and in organogenesis. Endogenous background β-galactosidase enzymatic activity was detectable at low levels from E14.5 onwards, but this was clearly distinguishable from reporter gene-specific expression (data not shown). At E11.5, β-gal was present at high levels in the surface ectoderm of Ecad-In2flox embryos (Fig. 4A). This expression was almost completely absent in corresponding Ecad-In2floxdel embryos (Fig. 4F), with only weak β-gal expression in the facial region, mandibulary and maxillary components of branchial arches, and AER. These differences were maintained at E12.5, when additional expression domains appeared in the follicles of vibrissae in Ecad-In2flox embryos (Fig. 4B) which were very weak in Ecad-In2floxdel embryos (Fig. 4G). During skin development between E14.5 to E16.5, the differences in β-gal activity between Ecad-In2flox and Ecad-In2floxdel became less pronounced, but were still obvious after only a brief incubation (45 minutes) in X-gal solution (compare Fig. 4C with 4H). Similarly, epithelia of the inner organs of E16.5 Ecad-In2floxdel embryos showed residual β-gal expression that was much weaker than that in Ecad-In2flox embryos (compare Fig. 4D,E with 4I,J, respectively).

When E11.5 Ecad-In2flox and Ecad-In2floxdel embryos were sectioned, high levels of β-gal expression were detected in the lens and ectoderm surrounding the eye of Ecad-In2flox embryos, whereas expression in this area was absent in Ecad-In2floxdel embryos (Fig. 4K,R). Likewise and in contrast to Ecad-In2flox embryos (Fig. 4L-Q), β-gal expression was not found in the surface ectoderm of the back (Fig. 4S), nasal cavity (Fig. 4T), stomach (Fig. 4U), gut (Fig. 4W) or metanephros (Fig. 4X) of Ecad-In2floxdel embryos. Only faint β-gal expression was observed in the pancreas primordium of Ecad-In2floxdel embryos (Fig. 4V). Collectively, these results suggest that the E-cadherin locus can be activated in later stages of development in a tissue-specific manner, even without the cis-regulatory elements of intron 2, but to a very reduced extent. During this later phase of development, sequences of intron 2 strongly enhance transcription of the E-cadherin reporter gene.

## Intron 2 sequences are not required for the E-cadherin reporter gene expression in the yolk sac

The results described above revealed that the presence of intron 2 had a more global enhancing effect on activation of E-cadherin transcription, particularly in later stages of development. During this analysis it became apparent that the β-gal expression in the yolk sac was independent of intron 2 sequences. Whereas yolk sacs of wild-type embryos do not show endogenous β-galactosidase expression at E10.5 (Fig. 5A) and only faint staining was observed at E12.5 (Fig. 5C), the yolk sacs of Ecad-In2flox and Ecad-In2floxdel embryos showed high-level reporter gene-derived β-gal expression (Fig. 5B,D). Remarkably, β-gal expression was equally high in the

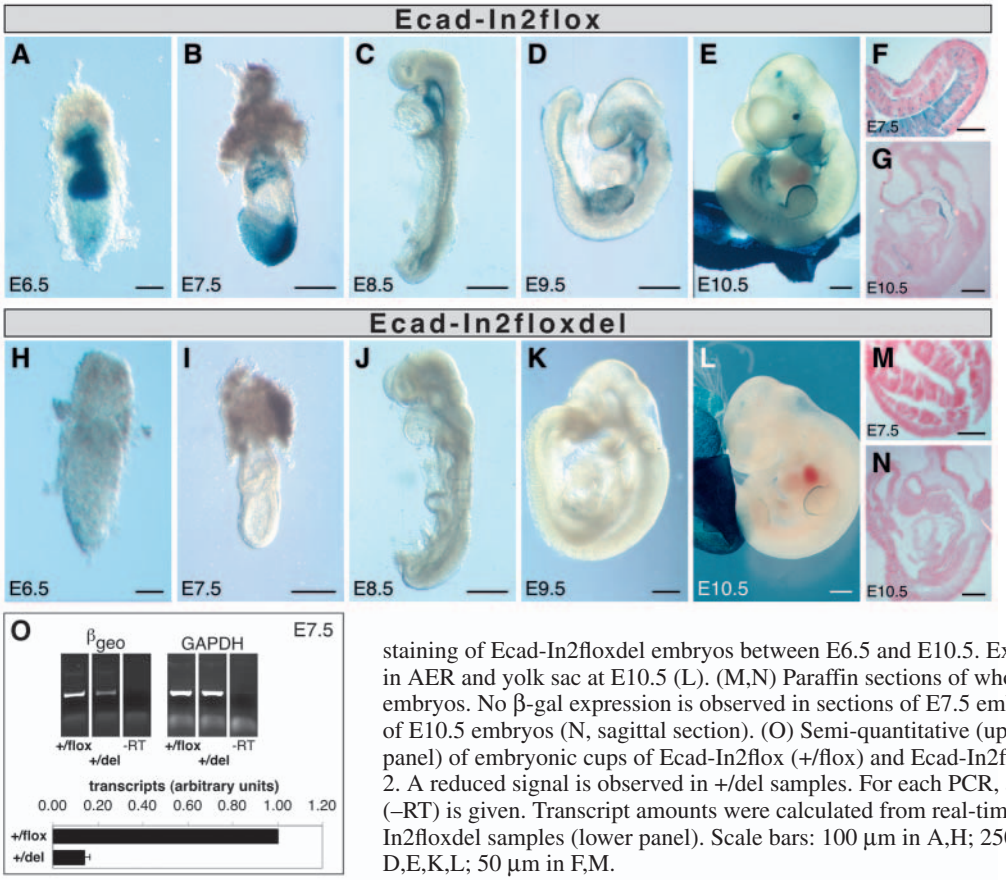


**Fig. 3.** Deletion of intron 2 leads to loss of β-gal expression during early embryogenesis. (A-E) Whole-mount X-gal staining of Ecad-In2flox embryos between E6.5 and E10.5. (F,G) Paraffin sections of whole-mount stained Ecad-In2flox embryos. Transverse section at E7.5 (F) indicates expression in ectoderm and endoderm, but not in mesoderm. Expression levels appear higher in posterior ectoderm. Sagittal section at E10.5 (G) shows high-level β-gal expression in pharynx and gut epithelium. Low-level expression in surface ectoderm is not visible in sections at this stage. (H-L) Whole-mount X-gal staining of Ecad-In2floxdel embryos between E6.5 and E10.5. Expression is only observed at low levels in AER and yolk sac at E10.5 (L). (M,N) Paraffin sections of whole-mount stained Ecad-In2floxdel embryos. No β-gal expression is observed in sections of E7.5 embryos (M, transverse section) or inside of E10.5 embryos (N, sagittal section). (O) Semi-quantitative (upper panel) and real-time PCR (lower panel) of embryonic cups of Ecad-In2flox (+/flox) and Ecad-In2floxdel (+/del) embryos, similar to Fig. 2. A reduced signal is observed in +/del samples. For each PCR, a control without reverse transcriptase (–RT) is given. Transcript amounts were calculated from real-time PCR to be 85% reduced in Ecad-In2floxdel samples (lower panel). Scale bars: 100 μm in A,H; 250 μm in B,C,G,I,J,N; 500 μm in D,E,K,L; 50 μm in F,M.

yolk sacs of both genotypes, although a clear difference was observed between the respective embryos (Fig. 5B,D). Semi-quantitative and real-time PCR corroborated the X-gal staining data showing intron 2-independent expression of β-gal in yolk sacs at E10.5 and E16.5 (Fig. 5E).

### E-cadherin reporter gene activity in the lens and salivary gland epithelium is completely dependent on intron 2 sequences

In contrast to the situation in yolk sac tissue, two different E-cadherin expression domains were identified where reporter gene expression was never detected in Ecad-In2floxdel even in late embryonic stages. Whereas intense blue X-gal staining was observed in the lenses of Ecad-In2flox embryos at E10.5 (Fig. 6A), E12.5 (Fig. 6C) and E14.5 (Fig. 6E), expression was absent in the lenses of Ecad-In2floxdel embryos (Fig. 6B,D,F). Similarly, in contrast to Ecad-In2flox (Fig. 6G,I, arrow), no β-gal expression was observed in salivary glands of Ecad-In2floxdel embryos at E16.5 (Fig. 6H,M, arrow), whereas expression in skin, thyroid glands (Fig. 6H,N, open arrowhead) and meninges (Fig. 6H,P, arrowhead) is still detected because of weak activity of the reporter gene during later embryonic development. These results show different requirements of intron 2 sequences for E-cadherin-specific β-gal expression in different organs.

### Intron 2 sequences are necessary for initial activation of the locus and for maintenance of expression

To investigate whether intron 2 is required also for maintenance of transcription after the initial activation of the locus, we deleted these DNA sequences conditionally during later development, after epithelia had already formed and E-cadherin expression had been initiated. The deletion of intron 2 from established epithelia was performed with two different transgenic *Cre*-recombinase-expressing mouse strains: CK14-*Cre* and CK19-*Cre* (Hafner et al., 2004; Harada et al., 1999). In CK14-*Cre* mice, *Cre*-expression is controlled by the cytokeratin 14 promoter, which drives expression in the developing skin (Hafner et al., 2004; Wang et al., 1997). In CK19-*Cre* mice, *Cre*-expression is driven by the cytokeratin 19 locus (knock-in) in the trophectoderm, and, from E8.0 onwards, in the notochord, definitive gut endoderm and endoderm-derived epithelia. At later stages, CK19-*Cre* is also expressed in the surface ectoderm (Harada et al., 1999; Tamai et al., 2000).

Using CK14-*Cre* to recombine the Ecad-In2flox locus, no difference in β-gal expression between the Ecad-In2flox and Ecad-In2flox/CK14-*Cre* was detected before E12.5 (data not shown). At E12.5, a slight reduction in β-gal expression was observed in the surface ectoderm of Ecad-In2flox embryos carrying the CK14-*Cre* allele (Fig. 7A, right, +/Δ) when compared with CK14-*Cre* negative embryos (Fig. 7A, left, +/flox). This difference became more evident at E13.5 and E14.5 (Fig. 7B,C). Interestingly, β-gal expression persisted in the lens and the gut loops of Ecad-In2flox/CK14-Cre embryos (compare left/right Fig. 7B), because the CK14-*Cre* is not expressed in these tissues (Hafner et al., 2004; Wang et al., 1997). At E16.5, intense β-gal expression was visible in the skin of control embryos (Fig. 7D, left), but only faint staining

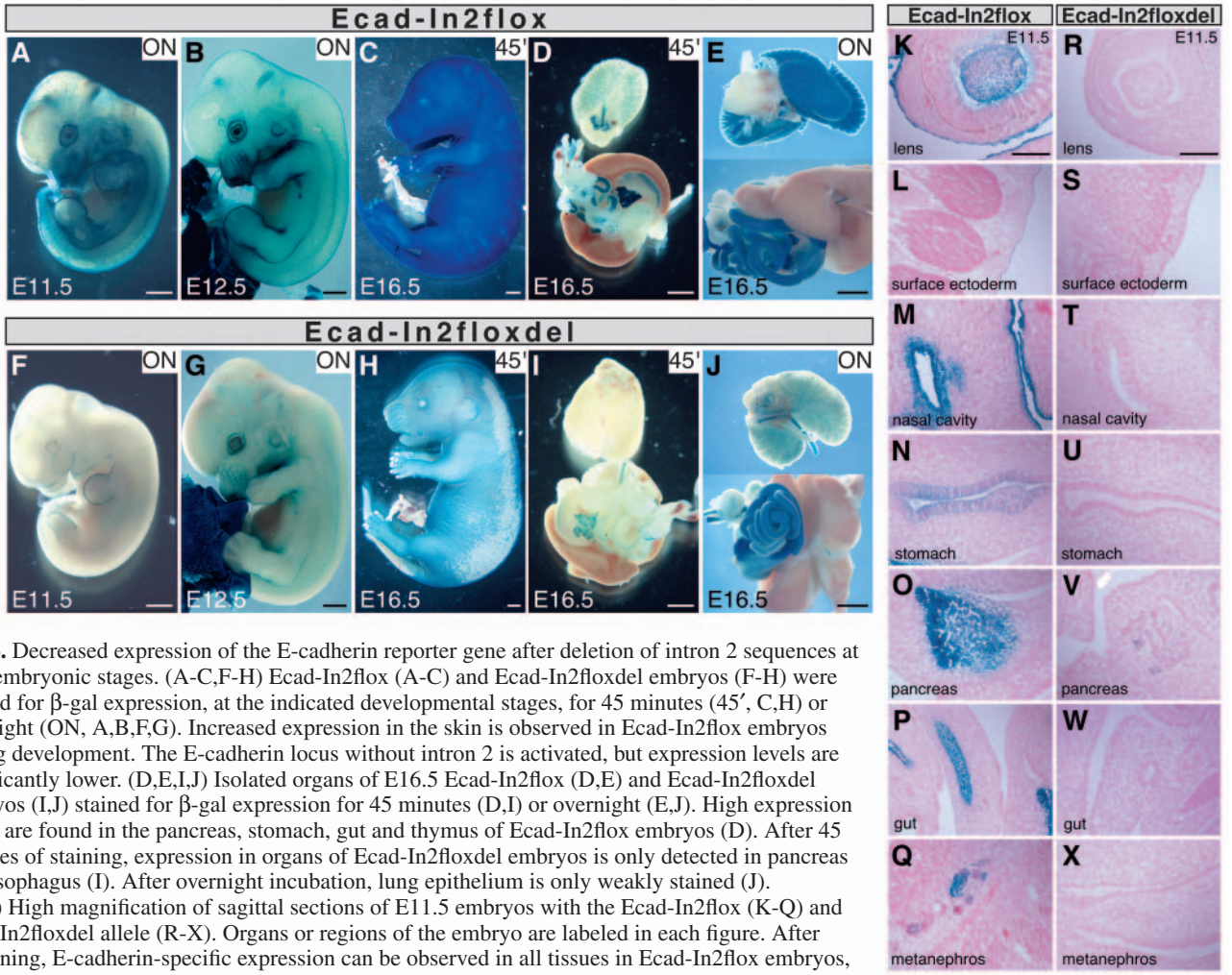


**Fig. 4.** Decreased expression of the E-cadherin reporter gene after deletion of intron 2 sequences at later embryonic stages. (A-C,F-H) Ecad-In2flox (A-C) and Ecad-In2floxdel embryos (F-H) were stained for β-gal expression, at the indicated developmental stages, for 45 minutes (45′, C,H) or overnight (ON, A,B,F,G). Increased expression in the skin is observed in Ecad-In2flox embryos during development. The E-cadherin locus without intron 2 is activated, but expression levels are significantly lower. (D,E,I,J) Isolated organs of E16.5 Ecad-In2flox (D,E) and Ecad-In2floxdel embryos (I,J) stained for β-gal expression for 45 minutes (D,I) or overnight (E,J). High expression levels are found in the pancreas, stomach, gut and thymus of Ecad-In2flox embryos (D). After 45 minutes of staining, expression in organs of Ecad-In2floxdel embryos is only detected in pancreas and esophagus (I). After overnight incubation, lung epithelium is only weakly stained (J). (K-X) High magnification of sagittal sections of E11.5 embryos with the Ecad-In2flox (K-Q) and Ecad-In2floxdel allele (R-X). Organs or regions of the embryo are labeled in each figure. After sectioning, E-cadherin-specific expression can be observed in all tissues in Ecad-In2flox embryos, but no expression is found after deletion of intron 2, except for a faint expression detected in the pancreas primordium (V). Scale bars: 1 mm in A-J; 100 μm in K-X.

was observed in the skin of Ecad-In2flox/CK14-*Cre* embryos (Fig. 7D, right; compare with Fig. 4D,I).

We obtained similar results when using CK19-*Cre* to ablate intron 2; at E9.5 and earlier, no difference in reporter gene activity was observed (Fig. 7E), but at E10.5, a significant reduction in β-gal expression level was observable in Ecad-In2flox/CK19-*Cre* embryos. Staining of the endoderm and also in the ectoderm was reduced compared with control embryos (Fig. 7F). This reduction was more obvious when examining sections of the stained embryos; β-gal expression in the gut tube was significantly reduced after CK19-*Cre* ablation of intron 2 (Fig. 7G,H). In the pharynx region, β-gal expression was mosaic following deletion of intron 2 sequences, presumably owing to incomplete deletion (Fig. 7I,J). These results indicate that intron 2 is required for maintaining E-cadherin gene activity in the gut epithelium and the skin, in addition to its role in initiation of transcription.

## Discussion

### E-cadherin transcriptional activity is faithfully recapitulated by the β-gal reporter allele

To monitor gene activity of the E-cadherin locus, we used the enzyme activity derived from the E-cadherin-*betageo* knock-in allele. In order to validate this approach, it was important to show that E-cadherin expression and β-gal activity coincide in a spatiotemporal manner. Both the Ecad-ATG (see Fig. S2 in the supplementary material) and the Ecad-In2flox knock-in alleles faithfully recapitulated all E-cadherin expression domains, and we did not observe any ectopic expression of the reporter gene. β-Gal activity was present as soon as zygotic E-cadherin expression is detected in four-cell stage embryos and was downregulated during gastrulation when mesodermal cells are formed. Thus, all changes in E-cadherin transcriptional activity are correctly reflected by β-gal activity.

### Complexity of E-cadherin transcriptional regulation

The position of cis-regulatory elements on genomic DNA sequences can be indicated by the presence of DNase-I-hypersensitive sites (DHSs). DHSs arise from nucleosome-free chromatin that is highly accessible to DNaseI and result from bound transcription factors. The occurrence of DHSs and the presence of cis-regulatory elements correlate in other genes (Harju et al., 2002; Kintscher et al., 2004; Lefevre et al., 2001; Murakami et al., 2004). At the E-cadherin locus, only one DHS is found upstream of the transcription start site at position

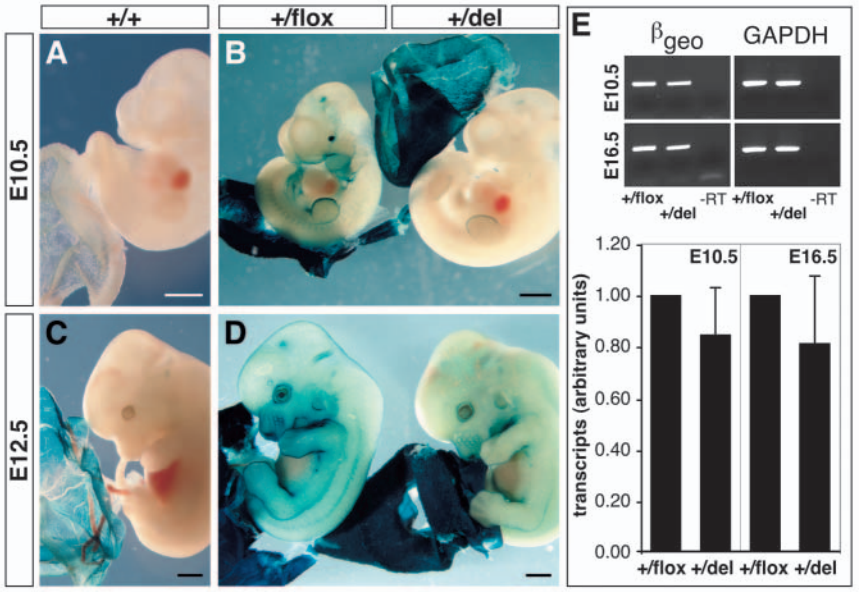


**Fig. 5.** E-cadherin-specific expression in the yolk sac is independent of cis-regulatory elements in intron 2. (A-D) Whole-mount β-gal staining of wild-type (A,C) or Ecad-In2flox (B,D, left) and Ecad-In2floxdel embryos and yolk sacs (B,D, right) at E10.5 (A,B) and E12.5 (C,D). Expression level in the yolk sac shows the same intensities in Ecad-In2flox and Ecad-In2floxdel embryos at all analyzed stages. (E) Semi-quantitative (upper panel) and real-time PCR (lower panel) of *betageo* (left) and *Gapd* transcripts (right) in yolk sacs at E10.5 and E16.5. No significant difference is observed in the level of expression between the two different strains. Scale bar: 1 mm.

–0.1 kb. The absence of additional DHSs further upstream of the E-cadherin promoter region and the accumulated occurrence of DHSs in the 5′ part of intron 2 in E-cadherin-expressing cells hint at cis-regulatory elements in intron 2 (Stemmler et al., 2003). A high degree of sequence conservation in mouse, rat, human, chimp and dog around these DNA elements in the part of intron 2 that has been analyzed for DHSs (–15 to +18 kb) further supports this notion (see Fig. S1 in the supplementary material). Because additional areas of significant sequence conservation in intron 2 outside of the DHS-mapped region were found, this suggested the existence of other regulatory elements spread over the entire intron 2.

In our previous transgenic reporter gene approach, we had demonstrated that a –6 to +0.1 kb promoter fragment is insufficient to drive E-cadherin expression. The first 15 kb of intron 2 sequences were beneficial towards properly regulating an E-cadherin transgene (Stemmler et al., 2003). This work also demonstrated that sequences required for E-cadherin-specific expression in the endoderm are found between +0.1 and +11 kb of the E-cadherin gene, a general enhancer between +11 and +16 kb, and a brain-specific enhancer between –6 and –1.5 kb. Nevertheless, it became evident that not all regulatory sequences have been covered by this analysis.

Nonetheless, encouraged by these findings and the fact that the entire intron 2 contained conserved sequences (see Fig. S1 in the supplementary material), we analyzed the function of these sequences in vivo by ablating the entire intron 2 using gene targeting. We were able to show that, if these sequences are deleted, E-cadherin expression is completely lost during early embryogenesis. Only during later embryonic development can the locus be activated without intron 2, but with significantly reduced expression levels. In addition, our analyses revealed even more complex regulatory functions of intron 2. In general, in expression domains that are affected by the absence of intron 2, these sequences are required for both activation of the locus and maintenance of expression. We found that, in the lens and the salivary glands, expression is absolutely controlled by cis-regulatory elements of intron 2, whereas, in the yolk sac, expression can be activated regardless of the presence of these sequences.

Based on our previous findings in transgenic mice (Stemmler et al., 2003) and the data presented here, we suggest the following model of regulating E-cadherin gene activity (Fig. 8). Whereas E-boxes at the promoter contribute to downregulating the locus (small red boxes, Fig. 8), E-cadherin gene activation is initiated and maintained due to intron 2 sequences. Importantly, in Ecad-In2floxdel embryos the endoderm-specific expression of the E-cadherin reporter gene was lost, probably owing to the lack of sequences between +1.2 and +11 kb (endoderm, Fig. 8). Entire loss of β-gal expression in Ecad-In2floxdel embryos can be partially ascribed to the general enhancer between +11 and +16 kb (enh., Fig. 8). However, ectoderm-specific expression is not at all detectable until E11.5 in Ecad-In2floxdel embryos nor was it consistently observed in the transgenic analysis (Stemmler et al., 2003). This indicates that additional, so far undescribed cis-regulatory elements in intron 2 are present between +18 and +47 kb to drive expression in the ectoderm (indicated by ‘tse’ in Fig. 8). E-cadherin-specific reporter gene expression in the brain due to the function of cis-regulatory elements between –6 and –1.5 kb (brain, Fig. 8) needs to be restricted to the E-cadherin expression domain by an as yet unknown brain-specific silencer (sil., Fig. 8), because we observed additional ectopic β-gal activity in the brain of transgenic embryos (Stemmler et al., 2003). Because this was not the case in Ecad-In2floxdel embryos, we conclude that the postulated brain-specific silencer must be located outside of intron 2.

A common mechanism of gene regulation of classical cadherins involving cis-regulatory elements in proximity to the transcription start site is suggested by the following observations. The genomic structure of classical cadherins exhibits a high degree of conservation between different species, as well as between different members of the cadherin family. They share a similar large second intron, and the promoter regions, e.g. those of E- and P-cadherin, have conserved DNA motifs (Faraldo and Cano, 1993). Additionally, it has been shown that the 5′-proximal promoters of other cadherins are insufficient to faithfully reflect endogenous expression. However, intron 2 is involved in proper gene function, for example of the chicken N-cadherin or the LCAM genes (Li et al., 1997; Sorkin et al., 1993). In addition to the local and promoter-proximal cis-regulatory elements provided by intron 2, correct expression of each member in the cadherin cluster may be achieved by the function of a higher order control element

over a greater distance. This is supported by the lack of significant sequence similarities in different species between the 3′ end of the P-cadherin gene and –6 kb of the E-cadherin gene. Because of this, the presence of additional and so far unconsidered cis-regulatory elements 5′ of the promoter is unlikely. A locus or general control region (LCR, Fig. 8) might exist at the cadherin cluster for proper expression of each member of the cluster, similar to the regulation of the *Hoxd* cluster or of *Mrf4* and *Myf5* (Fomin et al., 2004; Spitz et al., 2003). For the correct transcriptional control of the E-cadherin locus, the gene is then linked to this element via the proximal elements of intron 2 by factors that interact with the complex formed at the LCR. A similar mechanism can be postulated for classical cadherins outside of this cluster.

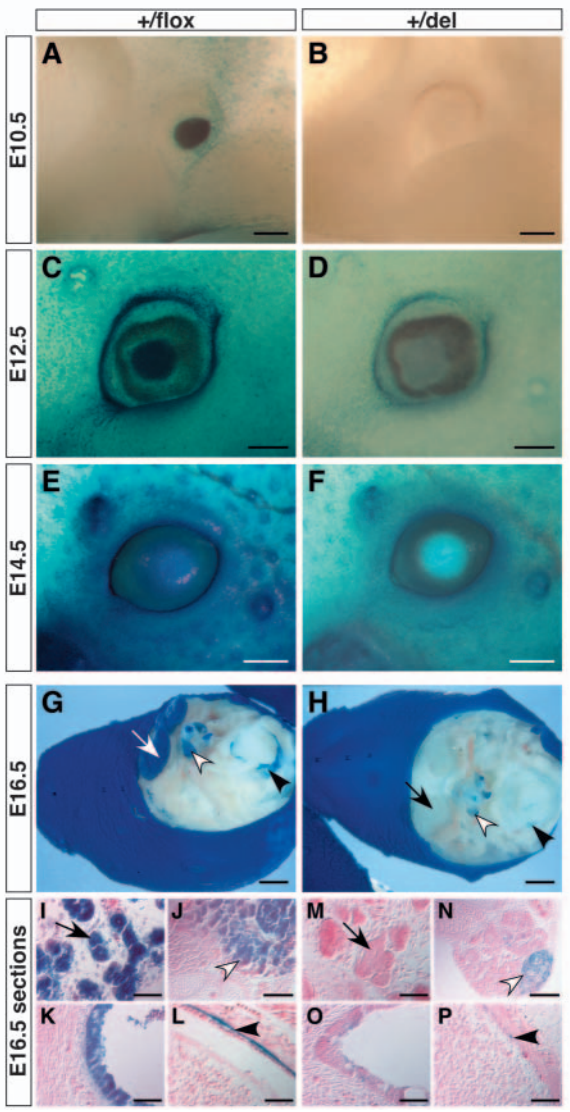


**Fig. 6.** Cells in the lens and salivary glands absolutely require cis-regulatory elements in intron 2 for activation of β-gal transcription. (A-F) X-gal staining in lenses of Ecad-In2flox embryos shows high intensities (A,C,E), whereas no expression is found in lenses of Ecad-In2floxdel embryos (B,D,F). Lenses at stages E10.5 (A,B), E12.5 (C,D) and E14.5 (E,F) are shown. (G-P) Analysis of E-cadherin-specific expression in mandibular salivary (arrow) and thyroid (white arrowhead) glands at E16.5 (black arrowhead marks meninges). Heads of Ecad-In2flox (G) and Ecad-In2floxdel embryos (H) were cut prior to X-gal staining and viewed from bottom. In Ecad-In2flox high expression is found in the salivary and thyroid glands (G), but no staining is observed in salivary glands of Ecad-In2floxdel embryos (H). (I-P) Transverse sections of Ecad-In2flox (I-L) and Ecad-In2floxdel heads at E16.5 (M-P). Sections of salivary (I,M) and thyroid glands (J,N), trachea (K,O) and meninges (L,P). Scale bars: 100 μm in A,B,J,N; 250 μm in C,D; 500 μm in E,F; 1 mm in G,H; 50 μm in I,K,L,M,O,P.

## Two mechanisms to initiate and maintain E-cadherin expression

We observed that, despite the lack of intron 2, the E-cadherin

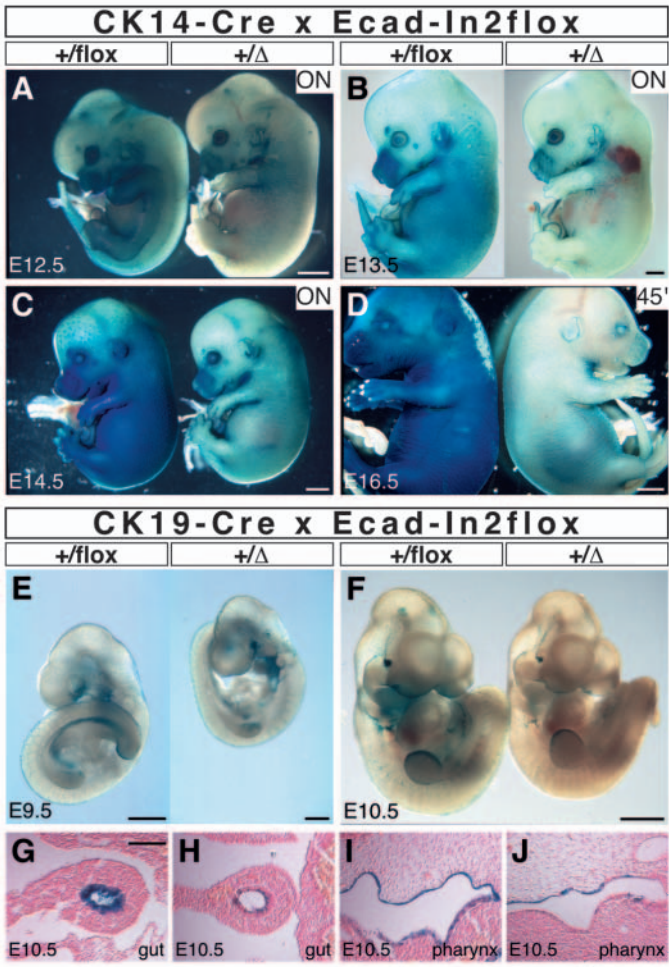


**Fig. 7.** Intron 2 sequences are required for maintaining E-cadherin expression. (A-D) Whole-mount β-gal staining of F1 embryos of CK14-*Cre* males crossed to Ecad-In2flox females. A slightly reduced staining is seen at E12.5 (A) in embryos where Cre was active (+/Δ) compared with control embryos with no *Cre* allele (+/flox). Further reduction is found in E13.5 (B), E14.5 (C) and E16.5 (D) embryos. Tissues where Cre was not active (lens, gut loops) are still strongly stained. (E,F) Whole-mount β-gal staining of F1 embryos of CK19-*Cre* males crossed to Ecad-In2flox females. No difference in gene activity is observed at E9.5 (E), but decreased gene activity after intron 2 deletion is visible in E10.5 embryos (F). (G-J) Transverse sections of the gut tube (G,H) and sagittal sections of the pharynx (I,J) of whole-mount stained E10.5 control (+/flox, G,I) and Ecad-In2flox/CK19-*Cre* embryos (+/Δ, H,J). Scale bars: 500 μm in E; 1 mm in A,B,F; 2 mm in C,D; 50 μm in G-J.

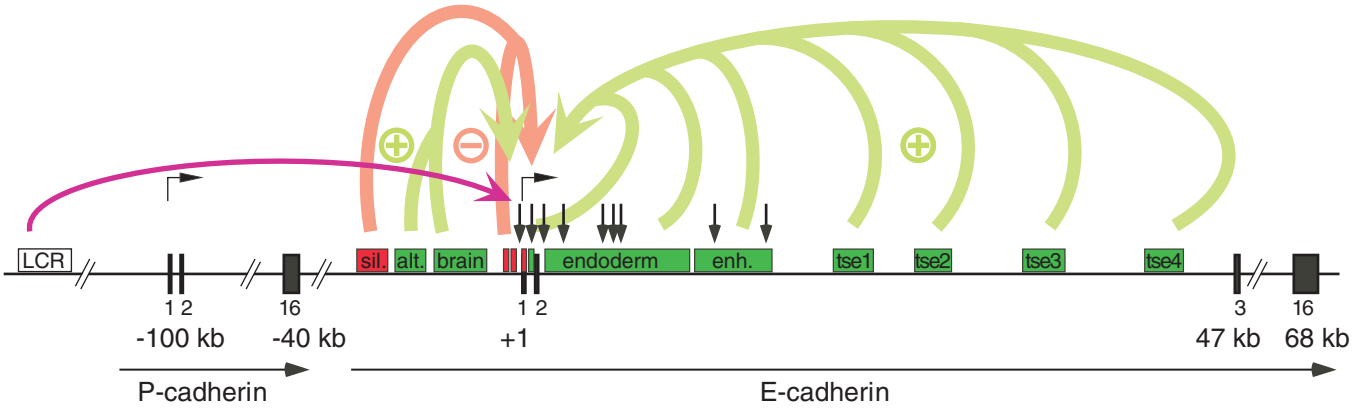


**Fig. 8.** Model of E-cadherin regulation in the cadherin cluster. Parts of the P- and E-cadherin locus are shown. Exons are represented by vertical black bars and numbers, DHSs by vertical arrows, transcription start sites by small horizontal arrows, E-boxes by red boxes and sequences with enhancing activities by green boxes (intron 1-enhancer represented by unlabeled box). alt., sequences that mediate alternative, intron 2-independent gene activation in late embryogenesis; brain, sequences that contribute to brain-specific expression; endoderm, sequences required for endoderm-specific expression; enh., sequences that generally enhance transcription; sil., brain-specific silencer that restricts gene activity to E-cadherin expression domains; tse1-4, tissue-specific enhancers, including elements for ectoderm-specific expression. The presence of a locus control region (LCR) is not yet proven and the precise positions of 'sil.', 'tse1-4' and 'alt.' are unknown. Elements that contribute to expression in yolk sac must be located outside of intron 2. E-boxes are required for downregulation of the gene (red arrow), whereas elements in intron 2 activate the locus (green arrows). The postulated LCR might influence gene activity for proper activation and downregulation (purple arrow).

locus was activated in many cell types of epithelial origin during late embryogenesis after E10.5. This suggests that the E-cadherin locus can be activated by two independent mechanisms. One mechanism acts during early embryogenesis and requires intron 2 for the onset of expression, and the second one functions at later stages. This second mechanism initiates E-cadherin expression independently of intron 2, although for high-level expression the support of the intron 2 enhancer elements is still required. The onset of the second wave of expression becomes apparent around E12.5 in the surface ectoderm (coinciding with the differentiation of the surface ectoderm and ongoing skin development) and in the gut endoderm. Presumably, this second, alternative activation mechanism is regulated by a common subset of transcription factors active in the specialized epithelia and might be achieved at the promoter or the intron 1 enhancer (Fig. 8).

### Different requirements of intron 2 sequences in certain specialized epithelia

Even more complexity of E-cadherin gene regulation emerged from the analysis of expression in the yolk sac, lens and salivary glands. The initiation of high-level reporter gene expression in the yolk sac is achieved independently of cis-regulatory elements of intron 2 and could reflect a gene-regulation mechanism specific to extra-embryonic tissues. By contrast, E-cadherin expression in the lens and the salivary glands is absolutely dependent on intron 2. Surprisingly, E-cadherin expression differs in tissues that originate from similar germ-layers. The lens develops from the lens placode, which is derived from surface ectoderm from E9.5 onwards. Whereas E-cadherin reporter gene expression is initiated by the second wave of expression in Ecad-In2floxdel embryos in the surface ectoderm of later stage embryos, no gene activation was found in the lens. Similarly, in epithelia of salivary glands of Ecad-In2floxdel embryos β-gal was never expressed, although they share the same germ-layer origin with epithelia of other inner organs. The postulated factors that are able to initiate E-cadherin transcription in later embryogenesis without intron 2 do not seem to be present in epithelia of salivary glands or in the lens. To explain the intron 2-dependent and independent E-cadherin expression, we propose that different tissue-specific enhancers probably exist that mediate E-cadherin expression in the yolk sac or in the lens and the salivary glands. This difference probably coincides with the different functions of specialized epithelia.

### The role of intron 2 in tumor progression

The data presented here reveal and emphasize the pivotal role of intron 2 in E-cadherin gene regulation during embryonic development. The importance of intron 2 sequences in gene regulation may also have an impact on tumorigenesis. The invasive property of cancer cells is often linked to loss of E-cadherin expression, in several cases owing to transcriptional downregulation (Berx et al., 1998). Accordingly, dysregulated expression of E-cadherin may be linked to mutations in intron 2 in cancer cells in which no mutation in the promoter or the coding sequence and no activation of a transcriptional repressor could be described. In some tumor cell lines, CpG-hypermethylation of the E-cadherin gene was discovered, but no mutation was found that might be responsible for this epigenic inactivation of the locus (Berx et al., 1998; Yoshiura et al., 1995). The mutations that are responsible for E-cadherin downregulation and subsequent CpG-hypermethylation may be located in intron 2. The identification of intron 2 mutations would underline the role of intron 2 in gene regulation in tumorigenesis. To be able to assess the impact of such mutations, a more precise description of the location and architecture of regulatory elements in intron 2 is required. Further gene targeting or transgenic mouse studies will concentrate on locating single tissue-specific cis-regulatory elements. An integrated in silico search for transcription factor binding sites can be used to determine which transcription

factors bind to the putative regulatory sequences of intron 2. Together, these approaches will lead to better understanding of the complex interplay of multiple regulatory regions dispersed throughout large parts of the E-cadherin locus.

We thank T. Krieg and C. Niessen for CK14-*Cre* mice, and M. Taketo for CK19-*Cre* mice. We are grateful to M. Leitges for providing E14.1 ES cells and sharing expert knowledge in gene targeting, and to B. Kanzler, E. Huber, L. Morawiec, N. Klemm and Y. Joos for support in the generation of transgenic mice. We thank K. Bruser and K. Hansen for excellent technical assistance, V. Taylor for helpful discussions, and R. Cassada for critically reading the manuscript. This work was supported by the GIF Research grant No. I-747-147.2/2002.

### Supplementary material

Supplementary material for this article is available at
http://dev.biologists.org/cgi/content/full/132/5/965/DC1

## References

**Batlle, E., Sancho, E., Franci, C., Dominguez, 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.

**Behrens, J., Lowrick, O., Klein-Hitpass, L. and Birchmeier, W.** (1991). The E-cadherin promoter: functional analysis of a GC-rich region and an epithelial cell-specific palindromic regulatory element. *Proc. Natl. Acad. Sci. USA* **88**, 11495-11499.

**Berx, G., Nollet, F. and van Roy, F.** (1998). Dysregulation of the E-cadherin/catenin complex by irreversible mutations in human carcinomas. *Cell Adhes. Commun.* **6**, 171-184.

**Bussemakers, M. J., Giroldi, L. A., van Bokhoven, A. and Schalken, J. A.** (1994). Transcriptional regulation of the human E-cadherin gene in human prostate cancer cell lines: characterization of the human E-cadherin gene promoter. *Biochem. Biophys. Res. Commun.* **203**, 1284-1290.

**Butz, S. and Larue, L.** (1995). Expression of catenins during mouse embryonic development and in adult tissues. *Cell Adhes. Commun.* **3**, 337-352.

**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.

**Carle, G. F., Frank, M. and Olson, M. V.** (1986). Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. *Science* **232**, 65-68.

**Carver, E. A., Jiang, R., Lan, Y., Oram, K. F. and Gridley, T.** (2001). The mouse Snail gene encodes a key regulator of the epithelial-mesenchymal transition. *Mol. Cell Biol.* **21**, 8184-8188.

**Chu, G., Vollrath, D. and Davis, R. W.** (1986). Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* **234**, 1582-1585.

**Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, 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.

**Conacci-Sorrell, M., Simcha, I., Ben-Yedidia, T., Blechman, J., Savagner, P. and Ben-Ze'ev, A.** (2003). Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of beta-catenin signaling, Slug, and MAPK. *J. Cell Biol.* **163**, 847-857.

**Detrick, R. J., Dickey, D. and Kintner, C. R.** (1990). The effects of N-cadherin misexpression on morphogenesis in Xenopus embryos. *Neuron* **4**, 493-506.

**Dymecki, S. M.** (1996). Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc. Natl. Acad. Sci. USA* **93**, 6191-6196.

**Fannon, A. M. and Colman, D. R.** (1996). A model for central synaptic junctional complex formation based on the differential adhesive specificities of the cadherins. *Neuron* **17**, 423-434.

**Faraldo, M. L. and Cano, A.** (1993). The 5′ flanking sequences of the mouse P-cadherin gene. Homologies to 5′ sequences of the E-cadherin gene and identification of a first 215 base-pair intron. *J. Mol. Biol.* **231**, 935-941.

**Fomin, M., Nomokonova, N. and Arnold, H. H.** (2004). Identification of a critical control element directing expression of the muscle-specific transcription factor MRF4 in the mouse embryo. *Dev. Biol.* **272**, 498-509.

**Fujimori, T., Miyatani, S. and Takeichi, M.** (1990). Ectopic expression of N-cadherin perturbs histogenesis in Xenopus embryos. *Development* **110**, 97-104.

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

**Gu, H., Zou, Y. R. and Rajewsky, K.** (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73**, 1155-1164.

**Hafner, M., Wenk, J., Nenci, A., Pasparakis, M., Scharffetter-Kochanek, K., Smyth, N., Peters, T., Kess, D., Holtkotter, O., Shephard, P. et al.** (2004). Keratin 14 Cre transgenic mice authenticate keratin 14 as an oocyte-expressed protein. *Genesis* **38**, 176-181.

**Harada, N., Tamai, Y., Ishikawa, T., Sauer, B., Takaku, K., Oshima, M. and Taketo, M. M.** (1999). Intestinal polyposis in mice with a dominant stable mutation of the beta-catenin gene. *EMBO J.* **18**, 5931-5942.

**Hardy, M. H.** (1992). The secret life of the hair follicle. *Trends Genet.* **8**, 55-61.

**Harju, S., McQueen, K. J. and Peterson, K. R.** (2002). Chromatin structure and control of β-like globin gene switching. *Exp. Biol. Med.* **227**, 683-700.

**Hatta, K., Takagi, S., Fujisawa, H. and Takeichi, M.** (1987). Spatial and temporal expression pattern of N-cadherin cell adhesion molecules correlated with morphogenetic processes of chicken embryos. *Dev. Biol.* **120**, 215-227.

**Hennig, G., Behrens, J., Truss, M., Frisch, S., Reichmann, E. and Birchmeier, W.** (1995). Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo. *Oncogene* **11**, 475-484.

**Hennig, G., Lowrick, O., Birchmeier, W. and Behrens, J.** (1996). Mechanisms identified in the transcriptional control of epithelial gene expression. *J. Biol. Chem.* **271**, 595-602.

**Hirai, Y., Nose, A., Kobayashi, S. and Takeichi, M.** (1989). Expression and role of E- and P-cadherin adhesion molecules in embryonic histogenesis. I. Lung epithelial morphogenesis. *Development* **105**, 263-270.

**Hogan, B. L.** (1999). Morphogenesis. *Cell* **96**, 225-233.

**Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M.** (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**, 292-295.

**Huber, O., Bierkamp, C. and Kemler, R.** (1996a). Cadherins and catenins in development. *Curr. Opin. Cell Biol.* **8**, 685-691.

**Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B. G. and Kemler, R.** (1996b). Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. *Mech. Dev.* **59**, 3-10.

**Jamora, C., DasGupta, R., Kocieniewski, P. and Fuchs, E.** (2003). Links between signal transduction, transcription and adhesion in epithelial bud development. *Nature* **422**, 317-322.

**Kintscher, J., Yamkamon, V., Braas, D. and Klempnauer, K. H.** (2004). Identification of a Myb-responsive enhancer of the chicken C/EBPbeta gene. *Oncogene* **23**, 5807-5814.

**Kuhn, R., Rajewsky, K. and Muller, W.** (1991). Generation and analysis of interleukin-4 deficient mice. *Science* **254**, 707-710.

**Larue, L., Ohsugi, M., Hirchenhain, J. and Kemler, R.** (1994). E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proc. Natl. Acad. Sci. USA* **91**, 8263-8267.

**Lefevre, P., Kontaraki, J. and Bonifer, C.** (2001). Identification of factors mediating the developmental regulation of the early acting –3.9 kb chicken lysozyme enhancer element. *Nucleic Acids Res.* **29**, 4551-4560.

**Li, B., Paradies, N. E. and Brackenbury, R. W.** (1997). Isolation and characterization of the promoter region of the chicken N-cadherin gene. *Gene* **191**, 7-13.

**Murakami, R., Osano, K. and Ono, M.** (2004). DNase I hypersensitive sites and histone acetylation status in the chicken Ig-beta locus. *Gene* **337**, 121-129.

**Nose, A. and Takeichi, M.** (1986). A novel cadherin cell adhesion molecule: its expression patterns associated with implantation and organogenesis of mouse embryos. *J. Cell Biol.* **103**, 2649-2658.

**Peinado, H., Marin, F., Cubillo, E., Stark, H. J., Fusenig, N., Nieto, M. A. and Cano, A.** (2004). Snail and E47 repressors of E-cadherin induce distinct invasive and angiogenic properties in vivo. *J. Cell Sci.* **117**, 2827-2839.

**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.

**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.

**Riethmacher, D., Brinkmann, V. and Birchmeier, C.** (1995). A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development. *Proc. Natl. Acad. Sci. USA* **92**, 855-859.

**Ringwald, M., Baribault, H., Schmidt, C. and Kemler, R.** (1991). The structure of the gene coding for the mouse cell adhesion molecule uvomorulin. *Nucleic Acids Res.* **19**, 6533-6539.

**Sambrook, J., Fritsch, E. F. and Maniatis, T.** (1989). *Molecular Cloning – A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

**Schwartz, D. C. and Cantor, C. R.** (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**, 67-75.

**Schwenk, F., Baron, U. and Rajewsky, K.** (1995). A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23**, 5080-5081.

**Shimamura, K. and Takeichi, M.** (1992). Local and transient expression of E-cadherin involved in mouse embryonic brain morphogenesis. *Development* **116**, 1011-1019.

**Shimamura, K., Takahashi, T. and Takeichi, M.** (1992). E-cadherin expression in a particular subset of sensory neurons. *Dev. Biol.* **152**, 242-254.

**Sorkin, B. C., Jones, F. S., Cunningham, B. A. and Edelman, G. M.** (1993). Identification of the promoter and a transcriptional enhancer of the gene encoding L-CAM, a calcium-dependent cell adhesion molecule. *Proc. Natl. Acad. Sci. USA* **90**, 11356-11360.

**Spitz, F., Gonzalez, F. and Duboule, D.** (2003). A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113**, 405-417.

**Stemmler, M. P., Hecht, A., Kinzel, B. and Kemler, R.** (2003). Analysis of regulatory elements of E-cadherin with reporter gene constructs in transgenic mouse embryos. *Dev. Dyn.* **227**, 238-245.

**Takeichi, M.** (1988). The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. *Development* **102**, 639-655.

**Tamai, Y., Ishikawa, T., Bosl, M. R., Mori, M., Nozaki, M., Baribault, H., Oshima, R. G. and Taketo, M. M.** (2000). Cytokeratins 8 and 19 in the mouse placental development. *J. Cell Biol.* **151**, 563-572.

**Thiery, J. P.** (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2**, 442-454.

**Vestweber, D. and Kemler, R.** (1984). Rabbit antiserum against a purified surface glycoprotein decompacts mouse preimplantation embryos and reacts with specific adult tissues. *Exp. Cell Res.* **152**, 169-178.

**Vestweber, D., Kemler, R. and Ekblom, P.** (1985). Cell-adhesion molecule uvomorulin during kidney development. *Dev. Biol.* **112**, 213-221.

**Wang, X., Zinkel, S., Polonsky, K. and Fuchs, E.** (1997). Transgenic studies with a keratin promoter-driven growth hormone transgene: prospects for gene therapy. *Proc. Natl. Acad. Sci. USA* **94**, 219-226.

**Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W., Stott, D. and Allemann, R. K.** (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. *Genes Dev.* **5**, 2048-2059.

**Wilkinson, D. and Green, J.** (1990). *Postimplantation Mouse Embryos: A Practical Approach.* Oxford, UK: IRL Press.

**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.

**Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T. and Hirohashi, S.** (1995). Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc. Natl. Acad. Sci. USA* **92**, 7416-7419.