Mice exclusively expressing the short isoform of Smad2 develop normally and are viable and fertile

N. Ray Dunn,[1] Chad H. Koonce,[2] Dorian C. Anderson, Ayesha Islam,[2] Elizabeth K. Bikoff,[2] and Elizabeth J. Robertson[2,3]

Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

Smad2 and Smad3 are closely related effectors of TGFβ/Nodal/Activin-related signaling. *Smad3* mutant mice develop normally, whereas *Smad2* plays an essential role in patterning the embryonic axis and specification of definitive endoderm. Alternative splicing of *Smad2* exon 3 gives rise to two distinct protein isoforms. The short Smad2(Δexon3) isoform, unlike full-length Smad2, Smad2(FL), retains DNA-binding activity. Here, we show that *Smad2(FL)* and *Smad2(Δexon3)* are coexpressed throughout mouse development. Directed expression of either *Smad2(Δexon3)* or *Smad3*, but not *Smad2(FL)*, restores the ability of *Smad2*-deficient embryonic stem (ES) cells to contribute descendants to the definitive endoderm in wild-type host embryos. Mice engineered to exclusively express *Smad2(Δexon3)* correctly specify the anterior–posterior axis and definitive endoderm, and are viable and fertile. Moreover, introducing a human *Smad3* cDNA into the mouse *Smad2* locus similarly rescues anterior–posterior patterning and definitive endoderm formation and results in adult viability. Collectively, these results demonstrate that the short Smad2(Δexon3) isoform or Smad3, but not full-length Smad2, activates all essential target genes downstream of TGFβ-related ligands, including those regulated by Nodal.

[*Keywords*: Smad2; Smad3; alternative splicing; chimeras; definitive endoderm; redundancy]

Transforming growth factor β (TGFβ)-related ligands comprise one of the largest groups of cytokines encoded in the vertebrate genome (Venter et al. 2001). They govern fundamental cell fate decisions in the embryo and are intimately involved in the maintenance of tissue homeostasis throughout postnatal life (for review, see Whitman 1998; Massagué et al. 2000). At the cell surface, the concentration of available ligand is interpreted by a receptor complex containing type I and type II serine–threonine kinases. Ligand-dependent activation of the type I receptor kinase triggers phosphorylation of a family of intracellular effector proteins termed Smads. The activated receptor-regulated Smads (R-Smads) then oligomerize with a structurally related collaborator, Smad4, and this heteromeric complex moves to the nucleus to regulate the transcription of target genes.

Smad2 and Smad3 function as R-Smads downstream of prototypical TGFβ ligands, Nodal, Activin, and some growth and differentiation factors (GDF) (Wall et al. 2000; Miyazawa et al. 2002; Cheng et al. 2003; Shi and Massagué 2003). These two molecules share 92% amino acid identity overall and display even greater similarity in the C-terminal MH2 protein–protein interaction domain (Fig. 1B). Flies and worms, unlike vertebrates, have only one *Smad2/3*-related gene, *dSmad2/Smox* and *sma-2*, respectively. It appears that gene duplication events within the chordate phyla gave rise to *Smad2* and *Smad3* (Newfeld et al. 1999; Dehal et al. 2002). In mice and humans, these loci show striking similarities in their linkage and genomic organization, with a near one-to-one correspondence in exon size (Fig. 1A). However, the prevailing model holds that *Smad2* and *Smad3* have significantly diverged and are nonredundant.

Recent work demonstrates that Smad2 and Smad3 exist in distinct oligomeric complexes at steady state. Thus, Smad2 is found mostly as a monomer within the cytoplasm, while Smad3 constitutively forms higher-order complexes (Jayaraman and Massagué 2000). After ligand-dependent phosphorylation, Smad2 and Smad3 are released from cytoplasmic retention proteins such as SARA and assemble into multimeric complexes with Smad4, and this activated complex is next imported into
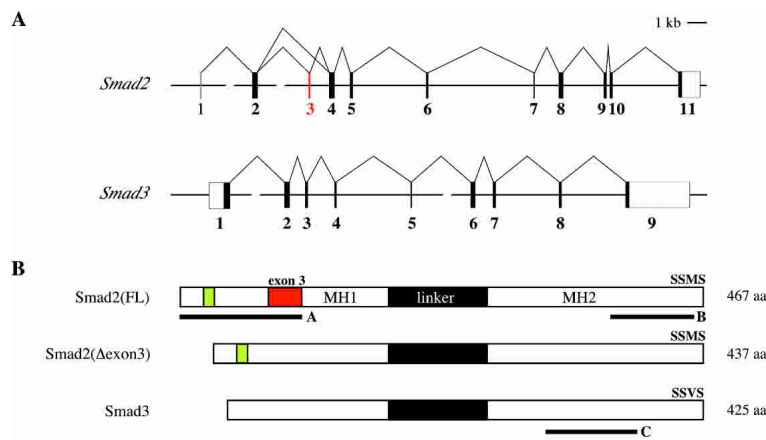
Figure 1. *Smad2* and *Smad3* structural comparisons. (*A*) Mouse *Smad2* and *Smad3* genomic organization. The alternatively spliced *Smad2* exon 3 and the corresponding 30-amino acid insert it encodes are shown in red. Noncoding exons are depicted as open rectangles. *Smad2* exons 4–11 and *Smad3* exons 2–9 are nearly identical in size. (*B*) Smad2 and Smad3 protein alignments showing the highly conserved N-terminal and C-terminal Mad Homology domains 1 (MH1) and 2 (MH2), respectively, and the intervening proline-rich linker region. The Smad2 MH1 domain also contains an N-terminal glycine-rich 10-amino acid insert (yellow) that is not present in Smad3. Smad2 and Smad3 possess distinct C-terminal SS(V/M)S phosphorylation motifs. The sequences corresponding to those protected by RPA probes are indicated as black lines.

the nucleus (Tsukazaki et al. 1998; Di Guglielmo et al. 2003). Smad2 interacts directly with nucleopore components via its MH2 domain, whereas Smad3 uses the classical importin-β-dependent nuclear transport pathway (for review, see Reguly and Wrana 2003; Xu and Massagué 2004). Smad3 binds DNA with low affinity via the novel β-hairpin DNA-interaction motif within the MH1 domain (Shi et al. 1998). In contrast, Smad2 cannot bind DNA due to the presence of a unique 30-amino acid insert that lies just N-terminal to the β-hairpin (Dennler et al. 1999; Yagi et al. 1999). Interestingly, the inability of Smad2 to interact with itself or other proteins at steady state or to use the importin-β nuclear transport pathway has been attributed to the presence of this unique insert encoded by the alternatively spliced exon 3 (Fig. 1A; Jayaraman and Massagué 2000; Kurisaki et al. 2001).

Heteromeric Smad complexes associate with an array of tissue-specific transcription factors, coactivators, and corepressors (for review, see Massagué and Wotton 2000; Miyazawa et al. 2002; Liu 2003). Consistent with their different mechanisms of nuclear import and DNA binding, several distinct transcriptional partners have been identified that specifically interact with either Smad2 or Smad3. For example, Smad3 associates with HNF4 to transactivate the *apolipoprotein C-III* promoter, while Smad3–FoxO protein complexes directly regulate the *p21Cip1* growth inhibitory gene to achieve TGFβ-dependent cytostasis (Chou et al. 2003; Seoane et al. 2004). In addition, Smad2 acts in combination with the forkhead transcription factor FoxH1 (FAST) and Smad4 to activate the *goosecoid* promoter, while Smad3-containing complexes inhibit expression (Labbé et al. 1998). Taken together, these observations strongly suggest that Smad2 and Smad3 play unique roles downstream of TGFβ-related growth factors.

Smad2 and Smad3 are further distinguished by their loss-of-function phenotypes. *Smad3* mutant mice are viable and fertile (Zhu et al. 1998; Datto et al. 1999; Yang et al. 1999). In contrast, *Smad2* mutant embryos fail to form the specialized extraembryonic signaling center known as the anterior visceral endoderm (AVE) (Waldrip

et al. 1998; Heyer et al. 1999). Consequently, these embryos lack anterior–posterior (A-P) polarity and are highly disorganized by embryonic day 8.5 (E8.5). The dramatic phenotypic difference between *Smad2* and *Smad3* mutant mice results from the expression of *Smad2* and not *Smad3* in the AVE (Tremblay et al. 2000). Recent experiments with *Smad2*-deficient embryonic stem (ES) cells and studies of conditional gene inactivation within the early epiblast reveal a second, focal requirement for high levels of *Smad2* in the specification of definitive endoderm during gastrulation (Tremblay et al. 2000; Vincent et al. 2003). Similarly, loss of *Smad3* in the context of one wild-type *Smad2* allele results in a failure to specify anterior primitive streak derivatives during gastrulation (Dunn et al. 2004). Consequently, these mutant embryos develop anterior truncations that are identical to those observed in embryos with decreased levels of the Nodal ligand (Vincent et al. 2003). From this more refined genetic analysis, Smad3 emerges as an essential component of the Nodal signal transduction pathway (Liu et al. 2004). Biochemical studies also show that Smad2 and Smad3 associate with the activated Alk4 type I Nodal receptor (Lebrun et al. 1999). Furthermore, in transcriptional activation assays Smad3 can regulate the ASE autoregulatory enhancer element from the *Nodal* gene, previously characterized as Smad2–FoxH1 dependent (Osada et al. 2000; Saijoh et al. 2000; Dunn et al. 2004). These data therefore support an opposing model that Smad2 and Smad3 are to a large degree functionally interchangeable.

Here, we investigate potentially shared or unique functional activities of Smad2 and Smad3 in the embryo. We have previously shown that descendants of *Smad2*-deficient ES cells are greatly compromised in their ability to colonize the definitive endoderm in chimeric mouse embryos (Tremblay et al. 2000). In reconstitution experiments, the *Smad2(Δexon3)* splice variant and *Smad3*, but not full-length *Smad2*, can compensate for loss of *Smad2* and restore these activities. Remarkably, germline deletion of exon 3 from the *Smad2* locus results in 100% viability, indicating that expression of full-length *Smad2* is dispensable. Finally, replacing

Smad2 coding sequences with *Smad3* also results in a homozygous viable phenotype. These results demonstrate unequivocally that the short Smad2(Δexon3) isoform or Smad3 activate all essential TGFβ signaling pathways.

Results

Expression of Smad2 isoforms in the mouse

Alternatively spliced *Smad2* transcripts have been previously documented in various human cell lines (Takenoshita et al. 1998; Yagi et al. 1999), and in frogs both *Smad2(Δexon3)* and *Smad2(FL)* are coexpressed in unfertilized eggs and during development (Faure et al. 2000). To investigate the expression of *Smad2* splice variants in mice, we used a ribonuclease protection assay (RPA) probe spanning the ATG-containing exon 2 and alternatively spliced exon 3 that distinguishes *Smad2(FL)* and *Smad2(Δexon3)* transcripts (Fig. 1B). The shorter *Smad2(Δexon3)* transcript is strongly expressed in ES cells and mouse embryos at all stages examined as well as in adult tissues (Figs. 2A, 4D [below]). *Smad2(FL)* transcript levels are consistently higher than *Smad2(Δexon3)*, but the *FL:Δexon3* ratio decreases at later stages (Fig. 2A). By late gestation (E16.5) the ratio is nearly 1:1 (data not shown). In adults, *Smad2(FL)* is more abundant in spleen and thymus, while *Smad2Δexon3* is barely detectable in liver (Fig. 4D, below). Among the diverse tissues analyzed here, the *FL:Δexon3* ratio is never greater than 3.5:1. These RNAse protection assays were repeated several times analyzing independent RNA samples to ensure that the results were reproducible. Our findings contrast with previous RT–PCR results that estimate *FL:Δexon3* expression ratios to be 10:1 in human cell lines (Takenoshita et al. 1998; Yagi et al. 1999). Taken together, these results show that the

Figure 2. *Smad2* splice variants are coexpressed during early mouse embryonic development. (*A*) Quantitative analysis of *Smad2(FL)* and *Smad2(Δexon3)* expression levels in CCE ES cell and embryonic total RNA at various stages. The size of the fragments protected by RPA probe A (Fig. 1B) is indicated. The ratio of *Smad2(FL)* to *Smad2(Δexon3)* expression decreases with developmental age. The expression ratios indicated by the numbers at the bottoms of the lanes were calculated by scanning the gels in a PhosphorImager and measuring the amount of radioactivity in each band to determine the precise ratios. (*B*) Total *Smad2* transcript levels monitored with RPA probe B (Fig. 1B).

Smad2(Δexon3) splice form is dynamically regulated and represents a much higher proportion of transcripts than previously estimated by RT–PCR. One possibility is that alternative Smad2 isoforms play unique roles in the transduction of TGFβ-related signals in the mouse embryo and adult.

Reconstitution of Smad2 signaling in Smad2-deficient ES cells and colonization of definitive endoderm in chimeric mouse embryos

We previously engineered the $Smad2^{Robm1}$ loss-of-function mutation in which the ATG containing exon 2 is deleted (Waldrip et al. 1998). Two independent homozygous mutant ES cell lines KT11 and KT15 were subsequently derived that constitutively express the *Rosa26 lacZ* reporter gene (Tremblay et al. 2000). These lines efficiently colonize ectodermal and mesodermal lineages when injected into wild-type host blastocysts, but are greatly compromised in their ability to contribute to the definitive endoderm (Heyer et al. 1999; Tremblay et al. 2000).

To assess the unique or common activites of Smad2(FL) and Smad2(Δexon3) isoforms during development, we developed a strategy to restore Smad2 signaling in these *Smad2*-deficient ES cell lines in gain-of-function transfection experiments. Thus, full-length cDNAs encoding N-terminally Flag-tagged human (h) Smad2(FL), Smad2(Δexon3), or the closely related Smad3 proteins were inserted into a modified version of the pCAGGS vector, which directs robust expression in ES cells and early embryos and confers hygromycin resistance (Fig. 3A; Niwa et al. 1991; Yagi et al. 1999). Since the parental KT11 and KT15 ES cells lack endogenous Smad2(FL) and Smad2(Δexon3) proteins (Tremblay et al. 2000; Dunn et al. 2004), drug-resistant clones were initially analyzed by flow cytometry (FACS) with an anti-Smad2/3 antibody to evaluate hSmad protein levels (Fig. 3A). Smad2(FL) and Smad2(Δexon3) transfected clones display expression levels approximately equal to endogenous Smad2 in wild-type ES cells, while Smad3 transformants exhibit increased levels, which reflects the contribution of endogenous Smad3 (Fig. 3A). Expression of introduced human Smad proteins was also examined by Western blot analysis using the anti-Flag antibody (Fig. 3B). Finally, transfected ES cells were also treated with TGFβ1 or Activin and analyzed with an anti-phospho-Smad2 antibody to verify that hSmad proteins were efficiently phosphorylated by native receptor complexes in ES cells (data not shown). Consistent with findings by Parisi et al. (2003), we find that wild-type CCE ES cells express *ActRIIA* and *ActRIIB* type II receptors as well as the *Alk2*, *Alk4*, and *Alk5* type I receptors (L. Oxburgh, N.R. Dunn, and E.J. Robertson, unpubl.). Collectively these experiments demonstrate that Flag-tagged human Smad2(FL), Smad2(Δexon3), and Smad3 proteins stably expressed in KT11/KT15 ES cell clones potentially function as effectors downstream of Nodal signals.

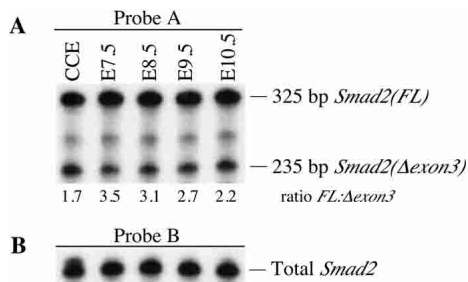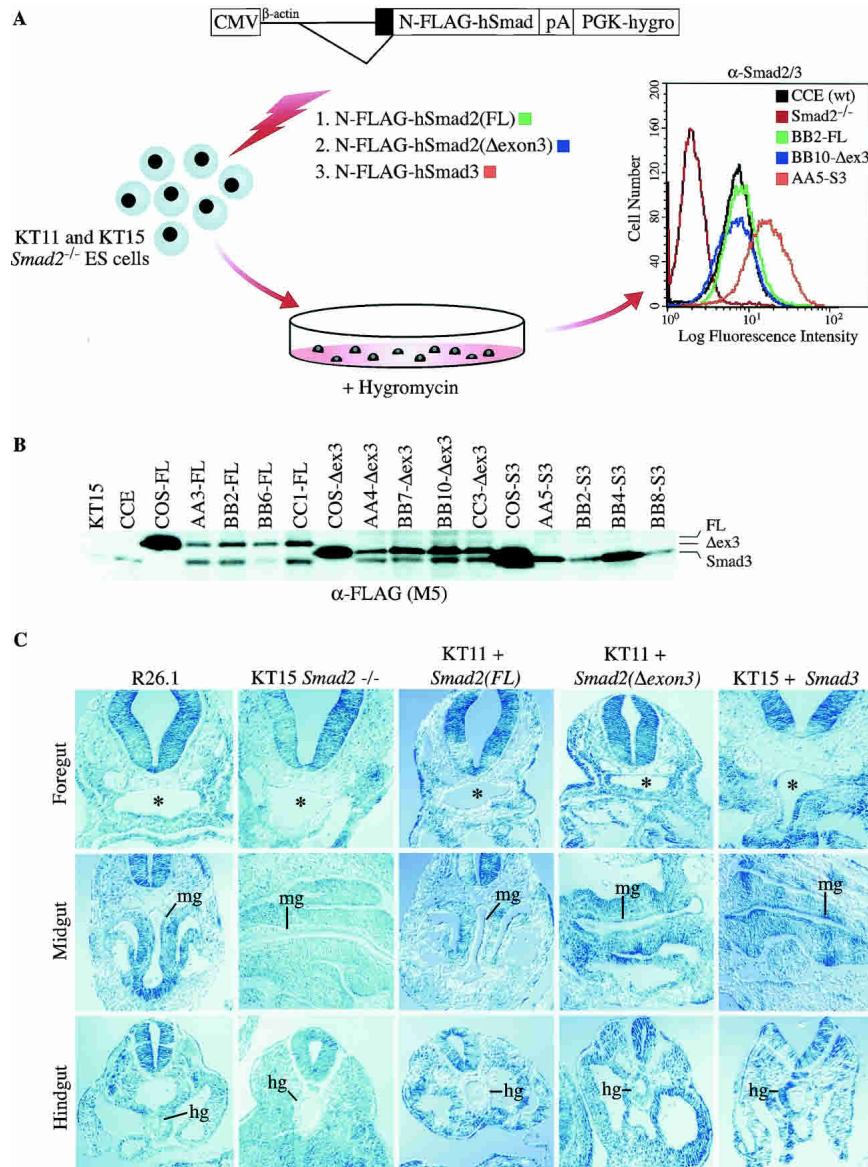Transformants were next tested for their ability to

Figure 3. Reconstituted *Smad2*-deficient ES cells expressing Smad2(Δexon3) or Smad3, but not Smad2(FL), contribute to the definitive endoderm in chimeric embryos. (*A*) Vector design and electroporation scheme. pCAGGS contains a chimeric promoter between the cytomegalovirus (CMV) immediate-early enhancer and the chicken β-actin promoter as well as the chicken β-globin intron and rabbit β-globin polyadenylation signals. cDNAs encoding N-terminally Flag-tagged human (h) Smad2(FL), Smad2(Δexon3), or Smad3 were introduced into *Smad2^Robm1* homozygous ES cells. Hygromycin-resistant clones were initially screened by flow cytometry for production of a human Smad2/3 protein. Smad2/3 expression levels in selected clones were subsequently analyzed via flow cytometry in experiments done synchronously with microinjection assays to verify that transfected ES cell clones tested in chimeras express roughly equivalent amounts of protein. A representative FACS is shown. (*B*) Western blot analysis reveals efficient production of Flag-hSmad2(FL), Flag-hSmad2(Δexon3), and Flag-hSmad3. The background band that comigrates with Flag-hSmad3 is detectable using the M5 monoclonal alone. (FL) Flag-hSmad2(FL); (Δex3) Flag-hSmad2(Δexon3); (S3) Flag-hSmad3. (*C*) Contribution to the definitive germ layers in E9.5 chimeric embryos. Transverse sections of representative chimeras stained for β-galactosidase activity at foregut (asterisk), midgut (mg), and hindgut (hg) levels. Wild-type R26.1 as well as KT11-Smad2(Δexon3)- and KT15-Smad3-expressing ES cells robustly colonize host embryos and give rise to descendants in all embryonic lineages, including the definitive endoderm. In contrast, Smad2-deficient KT15- and KT11-Smad2(FL)-expressing ES cells fail to contribute to the definitive endoderm, but efficiently form mesoderm and ectoderm.

contribute to the definitive endoderm in chimeric embryos (Table 1). As expected, wild-type *lacZ*-expressing R26.1 ES cells injected into wild-type host embryos efficiently colonize all embryonic germ layers, whereas *Smad2*-deficient KT15 ES cell descendants are specifically excluded from the definitive endoderm (Fig. 3C; Varlet et al. 1997; Tremblay et al. 2000). Transformants expressing hSmad2(Δexon3) or hSmad3 behave similarly to wild type, with ES cell derivatives (blue cells) interspersed among host wild-type cells along the length of the E9.5 gut epithelium (Table 1; Fig. 3C). It is particularly interesting that hSmad3 expression rescues Smad2 signaling because endogenous *Smad3* expression is not compromised in *Smad2*-deficient ES cells or embryos (Dunn et al. 2004). Thus, increased Smad3 levels appear

to compensate for the loss of Smad2 and permit gut colonization. Despite extensive chimerism throughout the embryo, none of the five KT11 subclones expressing hSmad2(FL) contribute to the definitive endoderm (Table 1; Fig. 3C). Thus, specification of definitive endoderm is rescued by expression of Smad2(Δexon3) or Smad3 in cells of the anterior primitive streak (APS), but Smad2(FL) expression on its own alone fails to restore Nodal-dependent signaling pathways.

Mutant mice exclusively expressing Smad2(Δexon3) *develop normally*

Results above strongly suggest that Smad2(Δexon3) but not Smad2(FL) plays an essential role during embryonic

Table 1. *Transfected ES cells expressing Smad2(Δexon3) or Smad3, but not Smad2(FL), contribute to the definitive endoderm in chimeric embryos*

Transfected ES cell clone	Chimeras analyzed	Percent colonization[a]	>10% β-gal positive cells in gut[a]
N-Flag-hSmad2(FL)[b]			
AA3	4	5%–60%	0
BB2	14	5%–80%	0
BB6	4	80%–90%	0
BB7	5	10%–90%	0
CC1	4	5%–70%	0
N-Flag-hSmad2(Δexon3)[b]			
BB2	6	5%–50%	1
BB10	9	30%–95%	8
CC1	1	20%	1
CC3	4	30%–90%	4
N-Flag-hSmad3[c]			
G5	16	30%–90%	14
H12	8	40%–90%	6

[a]Rescue indicates >10% of cells within the gut epithelium were derived from *lacZ*-expressing ES cell derivatives as assessed in serial sections of chimeric embryos.

[b]The parent ES cell line was KT11. We analyzed a total of 31 serially sectioned chimeric embryos obtained from five independent *hSmad2(FL)*-expressing clones. As for *Smad2*-deficient ES cells injected into host wild-type blastocysts, the *hSmad2(FL)*-expressing clones extensively contributed to the embryo, but their descendants were specifically excluded from the definitive endoderm.

[c]The parent ES cell line was KT15.

development. To further evaluate functional contributions made by the short form of Smad2, we engineered a novel *Smad2* allele lacking the alternatively spliced exon 3. *Smad2$^{\Delta exon3}$*/+ mice were derived and subsequently intercrossed to generate animals exclusively producing the *Smad2(Δexon3)* transcript (Fig. 3A–C). Surprisingly, *Smad2$^{\Delta exon3}$* homozygotes were born at Mendelian ratios, and matured to viable and fertile adults (data not shown). Adult homozygotes fail to display any overt late-onset abnormalities up to 9 mo of age.

To further characterize the *Smad2$^{\Delta exon3}$* mutant allele, total RNA was isolated from the organs of adult littermates and tested for expression of the *Smad2(FL)* and *Smad2(Δexon3)* transcripts. As expected, *Smad2$^{\Delta exon3}$* heterozygotes show increased levels of *Smad2(Δexon3)* compared to wild type, and *Smad2(FL)* expression is undetectable in *Smad2$^{\Delta exon3}$* homozygotes (Fig. 4D). Furthermore, loss of *Smad2(FL)* expression has no noticeable effect on endogenous *Smad3* expression (Fig. 4E; Dunn et al. 2004). We next assessed Smad2(FL) and Smad2(Δexon3) activation upon TGFβ1 stimulation. The anti-phospho-Smad2 antibody detects low levels of endogenous Smad2(Δexon3) protein in wild-type T lymphocytes, whereas this isoform is barely detectable in splenic B cells and HepG2 hepatoma cells (Fig. 4F). Stimulation with TGFβ1 leads to prominent Smad2(FL) phosphorylation. Despite roughly equal amounts of *Smad2(Δexon3)* mRNA expressed by embryonic and

adult tissues (Figs. 2A, 4D), the predominant species detectable by Western blot analysis is invariably Smad2(FL) (Fig. 4F), suggesting that *Smad2* variants are subject to significant post-transcriptional regulation.

We also evaluated Smad2(Δexon3) protein levels and function in the absence of Smad3. First, *Smad2$^{\Delta exon3}$*/+ and *Smad3^{null}*/+ mice were mated. *Smad2$^{\Delta exon3}$*; *Smad3^{null}* double heterozygous progeny were obtained in Mendelian numbers and then intercrossed. To our surprise, *Smad2$^{\Delta exon3/\Delta exon3}$*; *Smad3$^{null/null}$* pups were recovered at the predicted Mendelian frequency at weaning (data not shown). These animals are viable and fertile and display no overt hypomorphic phenotypes resulting from *Smad2$^{\Delta exon3}$*/*Smad3^{null}* genetic interactions, but do, however, develop osteoarthritis characteristic of *Smad3^{null}* homozygous adults (Datto et al. 1999; Yang et al. 2001; data not shown). We next characterized Smad2(Δexon3) protein levels and activation in *Smad2$^{\Delta exon3}$*; *Smad3^{null}* homozygous mutant mice. As expected, *Smad2$^{\Delta exon3}$*; *Smad3^{null}* homozygotes lack phospho-Smad2(FL), but Smad2(Δexon3) accumulates and is appropriately phosphorylated in response to TGFβ1 (Fig. 4F). Collectively, these results show that the Smad2(Δexon3) alone is fully capable of transducing the spectrum of TGFβ-related signals in the intact animal.

Human Smad3 *expressed under the control of the mouse* Smad2 *locus partially rescues viability*

Smad3 is not normally expressed within the visceral endoderm, a tissue in which *Smad2* is required to pattern the AVE. If Smad2(Δexon3) and Smad3 are functionally interchangeable, we reasoned that ectopic expression of Smad3 within the visceral endoderm should lead to normal establishment of embryonic A-P polarity and postnatal viability. To test this possibility, we designed a targeting vector to simultaneously delete coding sequences within exon 2 and introduce the N-terminally Flag-tagged *hSmad3* cDNA into the *Smad2* locus (Fig. 5A). To be able to distinguish the products arising from the manipulated and wild-type loci and to confirm that the introduced cDNA is appropriately expressed under control of the endogenous regulatory elements, and perhaps more importantly that there is no read-through expression of the endogenous wild-type product, we designed our knock-in alleles using Flag-tagged human Smads that were previously proven to produce biologically active proteins and developed a panel of RNAse protection probes that readily distinguish Flag-tagged human and endogenous mouse Smad transcripts. To avoid disruption of *cis*-acting regulatory elements that neighbor exon 2, we included the first 44 bp of exon 2 within the 5′ homology arm, while the 3′ homology arm overlaps with last 50 bp of exon 2. ES cells carrying the initial knock-in allele (*Smad2^{hSmad3kineo}*/+) were identified by Southern blot analysis. Derivative subclones lacking the neomycin selection cassette (*Smad2^{hSmad3ki}*/+) were recovered following transient transfection with a vector driving expression of Cre recombinase (Fig. 5A–C). *Smad2^{hSmad3ki}* heterozygotes were intercrossed to
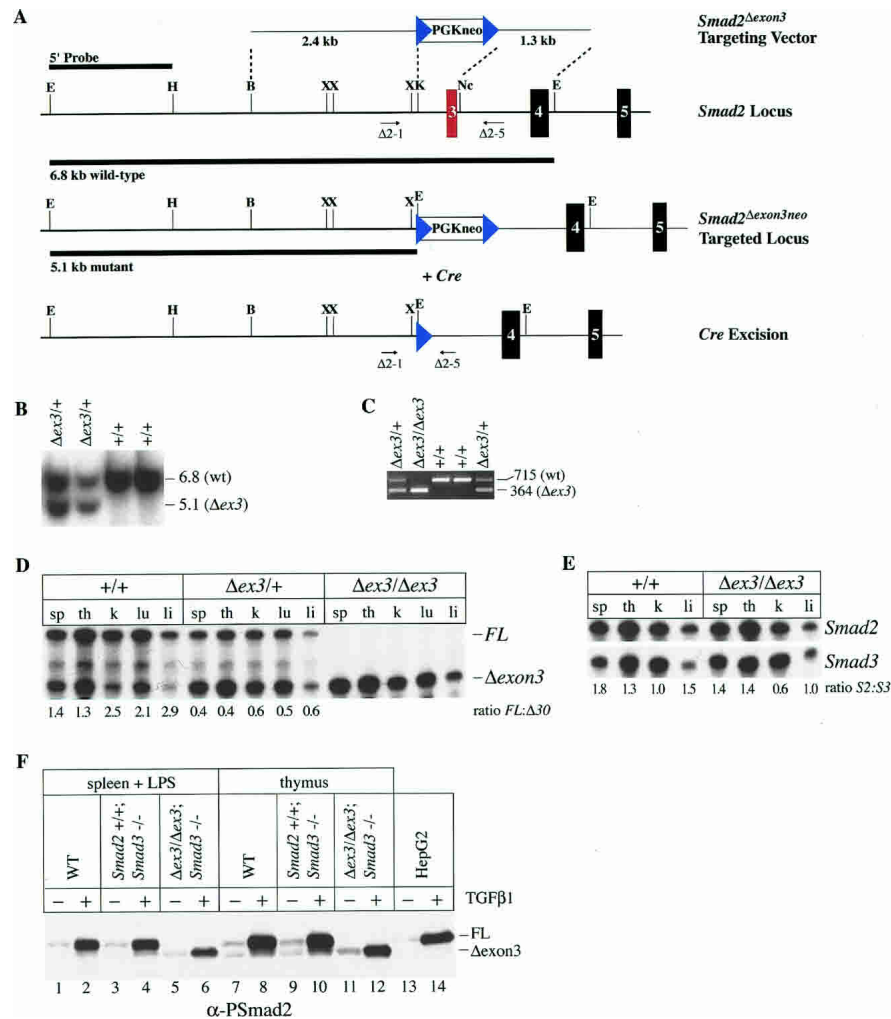
Figure 4. Production of mutant mice exclusively expressing the alternatively spliced Smad2Δexon3 isoform. (*A*) Strategy used to replace exon 3 with a *loxP* (blue triangles)-flanked *neomycin*-resistance cassette. Targeted clones were identified with a 5′ external probe (black line). (E) EcoRI; (H) HindIII; (B) BamHI; (X) XbaI; (K) KpnI; (Nc) NcoI. (*B*) Southern blot analysis. The 5′ external probe distinguishes 6.8-kb wild-type (+/+) and 5.1-kb targeted (Δex3/+) alleles. (*C*) PCR genotyping of wild-type, *Smad2*$^{\Delta exon3/+}$ (Δex3/+), and *Smad2*$^{\Delta exon3/\Delta exon3}$ (Δex3/Δex3) adult mice. The PCR primers in *A* give rise to products at the indicated sizes. (*D*) Total RNA from adult tissues was analyzed with RPA probe A (Fig. 1B), which distinguishes *Smad2(FL)* and *Smad2(Δexon3)* transcripts. *Smad2 (Δexon3)* mRNA expression is up-regulated in *Smad2*$^{\Delta exon3}$/+ mice, whereas *Smad2(FL)* transcripts are undetectable in *Smad2*$^{\Delta exon3}$ homozygous mutants. (*E*) Endogenous *Smad2* (probe B) and *Smad3* (probe C) expression levels are unperturbed in *Smad2*$^{\Delta exon3}$ homozygous mice. The expression ratios indicated by the numbers at the *bottoms* of the lanes were calculated by scanning the gels in a PhosphorImager and measuring the amount of radioactivity in each band to determine the precise ratios. (*F*) Western blot analysis using anti-phospho-Smad2 (α-PSmad2) reveals robust Smad2(FL) phosphorylation in wild-type B and T lymphocytes and HepG2 hepatocytes stimulated with TGFβ1, whereas Smad2(Δexon3) is efficiently phosphorylated in *Smad2*$^{\Delta exon3}$ homozygous mice (lanes *6,12*). (sp) Spleen; (th) thymus; (k) kidney; (lu) lung; (li) liver.

generate *Smad2*^{hSmad3ki} homozygous mice. Homozygotes were recovered at a sub-Mendelian frequency (10% vs. expected 25%) at weaning (Table 2), but those that survive to adulthood are viable and fertile.

To determine whether introduction of the *hSmad3* cDNA cassette and accompanying pA signal dysregulates transcription at the *Smad2*^{hSmad3ki} locus, we designed RPA probes that distinguish mouse and human *Smad3*. As expected, a 3′-probe derived from the *Flag-hSmad3* cDNA specifically detects *hSmad3* transcripts produced by the *Smad2*^{hSmad3ki} allele (Fig. 5D). Introduc-

tion of the triply repeated SV40 poly(A) cassette into the *Smad2* locus greatly attenuates transcription of the remaining downstream *Smad2* exons. Thus, residual 3′ *Smad2* transcripts are barely detectable in *Smad2*$^{hSmad3ki/hSmad3ki}$ adult tissues (Fig. 5D). Importantly, in *Smad2*^{hSmad3ki}/+ ES cells and viable adults, *hSmad3* transcripts arising from the manipulated locus are expressed at levels approximately equal to endogenous levels of *Smad2* (Fig. 5D).

To investigate the onset of lethality, *Smad2*^{hSmad3ki} homozygous embryos were collected at various develop-

Figure 5. Replacement of endogenous *Smad3* coding sequences with *Smad3*. (*A*) The ATG-containing (arrow) exon 2 of *Smad2* was replaced with an N-terminally Flag-tagged human (h) *Smad3* cDNA followed by a poly-adenylation signal and *loxP* (blue triangle)-flanked *neomycin*-resistance cassette. Targeted clones were identified with a 5′ external probe (black line). (B) BamHI; (E) EcoRI; (H) HindIII. (*B*) Southern blot analysis. The 5′ external probe detects 11.5-kb wild-type (+/+) and 5.1-kb targeted (*3ki/+*) alleles. (*C*) PCR analysis of DNA samples from wild-type, *Smad2^hSmad3ki^/+* (*3ki/+*), and *Smad2^hSmad3ki/hSmad3ki^* (*3ki/3ki*) adult mice. PCR primer locations are indicated in *A*. (*D*) Expression of the introduced *hSmad3* cDNA and endogenous *Smad2* by RPA of ES cell, human HepG2 hepatoma, adult tissues, and E12.5 total RNA. (Lanes *3,4*) A 3′ human *Smad3* RPA probe fails to protect mouse mRNA sequences. Similarly, probe B (Fig. 1B) is specific for mouse *Smad2* (cf. lanes *2* and *1,3–8*). The *hSmad3* RPA probe reveals a polymorphism between the HepG2 *Smad3* cDNA and the human *Smad3* cDNA used in our studies (cf. lanes *1* and *2*). (Lanes *1,5–8*) Small differences in the ratio of human *Smad3* expression from the *Smad2^hSmad3ki^* allele to *Smad2* expression from the wild-type allele are observed in *Smad2^hSmad3ki^/+* ES cells and mice. (Lanes *9–11*) Transcription from the remaining downstream exons in the *Smad2^hSmad3ki^* targeted locus is nearly undetectable. The expression ratios indicated by the numbers at the *bottoms* of the lanes were calculated by scanning the gels in a PhosphorImager and measuring the amount

of radioactivity in each band to determine the precise ratios. (*E*) Whole-mount views of E12.5 *Smad2^hSmad3ki^* homozygous embryos (*3ki/3ki*) derived from a homozygous parental intercross. Embryos are ordered from overt wild-type morphology (*left*) to increasingly more severe anterior abnormalities (*right*).
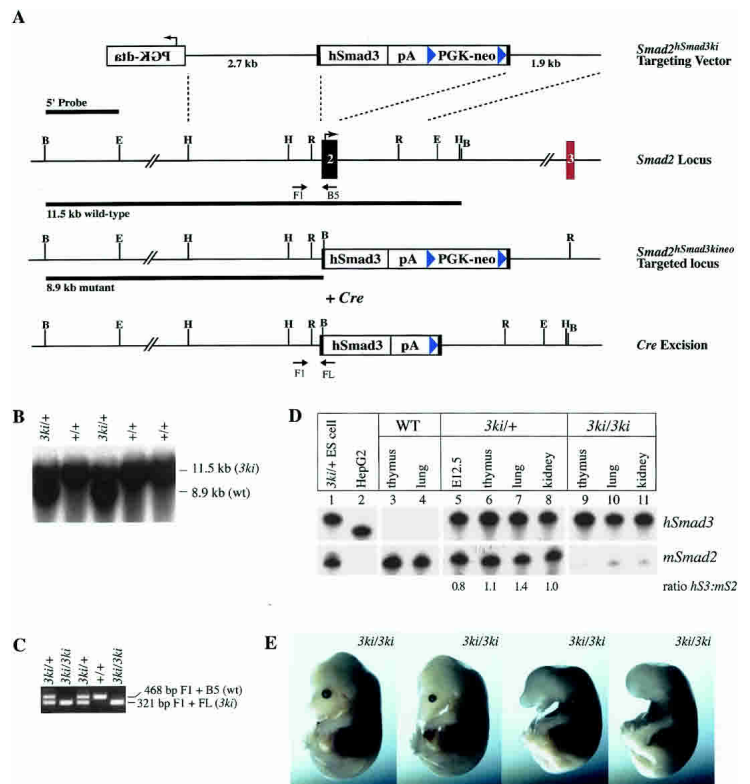
mental stages. From mid- to late gestation, abnormal embryos were easily recognized due to varying degrees of microcephaly, holoprosencephaly, or anencephaly (Fig. 5F). Histological analysis of six representative mutant embryos at E15.5 and E16.5 reveals limited forebrain and hindbrain development, with no associated defects in the situs of internal organs (data not shown). These anterior patterning defects closely resemble those arising from modulation of Nodal signaling within the APS (Nomura and Li 1998; Song et al. 1999).

One plausible explanation that accounts for the hypomorphic nature of the *Smad2^hSmad3ki^* allele is that the presence of the Flag epitope impacts the DNA binding or nuclear import functions within the MH1 domain, or alternatively augments the autoinhibitory interactions between the MH1 and MH2 domains (Hata et al. 1997). Additionally, Smad2 contains a 10-amino acid glycine-rich insert within the extreme N-terminal MH1 domain not found in Smad3 (Fig. 1B) that potentially regulates selection of DNA-binding partners. Finally, the intro-

Table 2. *Summary of phenotypic disturbances observed in* Smad2^hSmad3ki^ *intercross progeny*

Age	+/+	Smad2^hSmad3ki^/+	Smad2^hSmad3ki^/Smad2^hSmad3ki^	RS/ED[a]	AAD[b]
E9.5	15	16	6	4	0
E15.5	4	11	4	0	2
E16.5	15	28	8	6	4
Total (%)[c]	34 (32%)	55 (51%)	18 (17%)		
Weanlings (%)[c]	33 (32%)	60 (58%)	10 (10%)		

[a]The presence of resorption sites (RS) or empty decidua (ED) at the time of dissection suggests that some mutants die with more severe developmental abnormalities. Insufficient material for genotype determination.
[b]Number of *Smad2^hSmad3ki/hSmad3ki^* embryos with abnormal anterior development (AAD).
[c]Percentage of total embryos or weanlings genotyped. A loss of more than half of *Smad2^hSmad3ki^* homozygotes was observed across more than three generations of breeding. The surviving homozygotes show normal fertility, and in homozygous crosses, ~50% of the embryos develop normally, whereas the others die in utero or at birth due to anterior patterning defects. Since the *Smad2^hSmad3ki^* mice were maintained on a mixed (C57BL/6J × 129Sv//Ev)_F1 background, these distinct phenotypes may be due to genetic background differences.

duction of the human versus mouse *Smad3* cDNA by homologous recombination could lead to subtle functional differences, as has been recently illustrated with mice humanized for the EGF receptor (Sibilia et al. 2003). Nevertheless, the fact that a significant number of *Smad2^(hSmad3ki)* homozygotes are viable provides compelling evidence that Smad2(Δexon3) and Smad3 are the key regulators of TGFβ-signaling pathways.

Discussion

Smad2(FL) is nonessential for viability

Recent transcriptional activation assays comparing *Smad2*-deficient and *Smad3*-deficient mouse embryonic fibroblasts (MEFs) have identified distinct Smad2 and Smad3 downstream target genes. For example, *Smad7* expression requires Smad3, whereas *matrix metalloproteinase 2* regulation is dependent on Smad2 (Piek et al. 2001). Moreover, transcriptional profiling of these mutant MEFs reveals that in response to TGFβ stimulation, Smad3 activates a set of immediate-early genes that encode signal transducers and transcriptional regulators, while Smad2 appears to negatively modulate a number of these same genes (Yang et al. 2003). Similarly, morpholino knock-down experiments in HaCaT keratinocytes demonstrate that Smad3 serves as the principal effector of TGFβ-mediated growth arrest, while loss of Smad2 has no detectable effect (Kretschmer et al. 2003). Thus, considerable evidence suggests that the closely related Smad2 and Smad3 effectors are not redundant but, rather, have distinct activities and transcriptional targets. This difference is likely due to the presence of the unique 30-amino acid insert found just N-terminal to the β-hairpin in the Smad2 MH1 domain.

Here, we find quite unexpectedly that replacing the mouse *Smad2* gene with a Flag-tagged human *Smad3* cDNA results in adult viability. Thus, *Smad3* expressed under the control of the endogenous *Smad2* locus correctly regulates the panoply of target genes that lies downstream of TGFβ-related signals and is required for patterning the early embryo. This result is especially surprising since some of these targets such as *Nodal* and *Pitx2* were previously characterized by in vitro assays to be Smad2-dependent (Saijoh et al. 2000; Shiratori et al. 2001). In particular, we have shown that Smad2 is absolutely required in the visceral endoderm to specify the AVE (Waldrip et al. 1998). *Smad3* is not normally expressed in this tissue. The ectopic expression of *Smad3* by the *Smad2^(hSmad3ki)* allele restores Smad2 signaling and unequivocally demonstrates that Smad3 can activate the genetic program that specifies the AVE signaling center. Thus, irrespective of well-documented differences in DNA binding, nuclear import, oligomeric states, and interacting partners, the present experiments provide clear evidence that Smad2 and Smad3 are functionally interchangeable in the intact animal. In particular, it is the *Smad2(Δexon3)* transcript generated by alternative splicing that contributes essential activities, whereas Smad2(FL) is nonessential.

Smad2 *and* Smad3 *genetic interactions revisited*

Smad2^(+/−); Smad3^(−/−) mouse embryos display anterior patterning defects identical to conditional loss of *Smad2* within the epiblast (Vincent et al. 2003; Dunn et al. 2004). These strikingly similar phenotypes result from the loss of anterior definitive endoderm and prechordal plate that emerge from the APS during early gastrulation, and suggest that Smad2 serves as the predominant intracellular effector of Nodal signaling in the early embryo (Dunn et al. 2004). In other words, Smad3 levels within the epiblast of *Smad2* conditional mutant embryos as well as in chimeras mostly composed of *Smad2*-deficient ES cell descendants are inadequate to compensate for the loss of *Smad2*. We were unable to precisely determine the expression ratios of *Smad2(FL)*, *Smad2(Δexon3)*, and *Smad3* within the small population of epiblast cells that ingresses through the APS, but our results nevertheless demonstrate that the Smad2(Δexon3) isoform acting alone can correctly specify APS derivatives that come to underlie and pattern the anterior neural plate. Our results therefore suggest that the failure to pattern the APS in these mutant contexts results from the specific loss (or decreased levels) of *Smad2(Δexon3)*.

Several elegant studies with Activin and TGFβ1 have shown that the number of occupied cell surface receptors is proportional to the nuclear concentration of activated Smad2 (Shimizu and Gurdon 1999; Bourillot et al. 2002; Inman et al. 2002). Decreased concentrations of intracellular Smads are therefore predicted to significantly impact target gene regulation. Consistent with this, our previous work shows that the population of cells most sensitive to fluctuating Nodal levels resides within the APS (Norris et al. 2002; Vincent et al. 2003; Dunn et al. 2004). In the wild-type embryo, coexpression of the short Smad2(Δexon3) isoform and Smad3 may serve to selectively amplify Nodal signals in these discrete cell types.

We also engineered *Smad2^(hSmad2(FL))* and *Smad2^(hSmad2(Δexon3))* knock-in mutations via the identical strategy described above for the construction of the *Smad2^(hSmad3ki)* allele. Disappointingly, both these alleles recapitulated the *Smad2*-null phenotype, with characteristic failure to establish A-P polarity and formation of abundant extraembryonic mesoderm at E8.5. Homozygous mutant embryos were found to express Flag-hSmad2(FL) and Flag-hSmad2(Δexon3) proteins at levels equivalent to wild-type Smad2 (data not shown). In contrast, microinjection experiments analyzing transfected *Smad2*-deficient ES cells expressing Flag-tagged Smad2 isoforms for their ability to reconstitute chimeras exploited a strong ubiquitous promoter to drive expression throughout development, and in this case rescue of gut colonization only requires Smad2 activity over a narrow time window and at restricted tissue sites. In contrast, complete rescue of normal embryonic development requires appropriate expression of knock-in alleles under control of the endogenous regulatory elements governing wild-type *Smad2* levels at earlier egg cylinder stages in the visceral endoderm and the epiblast. The inability of Flag-tagged human Smad2 isoforms to rescue early em-

bryos is, indeed, quite perplexing. It seems most likely that then N-terminal Flag-epitope more strongly interferes with Smad2 than Smad3 functional activities. Nevertheless, the present positive results conclusively demonstrate that roughly half of the *Smad2[hSmad3ki]* homozygotes develop normally and are viable and fertile. Thus, we conclude that Smad3 can activate all essential TGFβ/Nodal/Activin signaling pathways.

Evolution of vertebrate Smad2 and Smad3

Nodal homologs have been identified in both chordates and echinoderms (Morokuma et al. 2002; Yu et al. 2002; Duboc et al. 2004). In sea urchins, *Nodal* expression in the presumptive oral ectoderm is absolutely required for the development of oral structures and regulates *Bmp2/4*, which acts downstream of Nodal to specify aboral fates (Duboc et al. 2004). This signaling pathway parallels the situation in mice in which *Bmp4* expression in the extraembryonic ectoderm is positively regulated by Nodal signals emanating from the proximal epiblast (Brennan et al. 2001; Beck et al. 2002). Thus, we propose that Nodal and its requisite intracellular effector Smad2 are components of an evolutionarily conserved signaling paradigm for organizing the basic body plan. Indeed, the sea urchin genome project has recently identified a *Smad2/3*-related gene (GenBank Trace Repository; http://sugp.caltech.edu).

In mice, *Smad2* is tightly linked to *Smad4* and *Smad7*, which encodes an inhibitory Smad, on chromosome 18, while *Smad3* is linked to *Smad6*, a second inhibitory Smad, on chromosome 9 (http://www.ensembl.org). This parallel genomic organization is also conserved in humans (OMIM). It therefore appears that *Smad3* arose in the vertebrate lineage due to a partial duplication of an ancestral "Smad" cassette. Since *Smad3*-null mutants exhibit phenotypes primarily confined to the regulation of the immune system, we therefore speculate that this duplication event coincides with the emergence of adaptive immunity among the jawed vertebrates (Datto et al. 1999; Yang et al. 1999; Kasahara et al. 2004). This prediction is also consistent with our dose-dependency model, which emphasizes the predominant role of Smad2 in axis specification and germ-layer formation in the early embryo, with amplifying or reinforcing activity provided by Smad3 in the primitive streak (Dunn et al. 2004). Indeed, a high percentage of chimeras derived from *Smad2*-deficient ES cells, whose definitive endoderm is wild type in origin, are overtly normal and fertile (our unpublished results).

Rare inactivating mutations in *Smad2* have been identified in a proportion of human colorectal and lung cancers, whereas similar genetic lesions in *Smad3* have not been characterized (Arai et al. 1998; Derynck et al. 2001; Siegel and Massagué 2003). It will be interesting to learn whether the tumor-suppressor activity of *Smad2* is contributed by the short Smad2(Δexon3) isoform. In a recent study, reduced Smad3 levels were observed in more than one-third of gastric tumors (Han et al. 2004). The loss of TGFβ-dependent cell cycle arrest and promotion of tu-

morigenesis may therefore be accomplished by the combined levels of Smad2(Δexon3) and Smad3 falling below a threshold that normally ensures the activation of the cyclin-dependent kinase inhibitors *p15* and *p21* and the maintenance of cytostasis (Feng et al. 2000; Pardali et al. 2000). Release from the antiproliferative effects of TGFβ may also occur by direct phosphorylation of Smad2(Δexon3)/Smad3 by the cyclin-dependent kinases, which inhibits their transcriptional activity (Matsuura et al. 2004).

Splicing among the Smad genes

To date, alternative splicing has been described for the *Smad2*, *Smad3*, *Smad4*, *Smad5*, and *Smad8* genes (Takenoshita et al. 1998; Nishita et al. 1999; Pierreux et al. 2000; Kjellman et al. 2004). For example, six *Smad4* splice variants have been identified that differ in the length of the linker region. Each of these isoforms retains its ability to interact with Smad2 and Smad3, but only three are capable of robust transcriptional activation (Pierreux et al. 2000). We speculate that structurally distinct Smad protein isoforms equip target cell populations with a versatile intracellular machinery to accommodate the array of extracellular TGFβ-related signals, and provide an important mechanism for modulating target gene activity (Lareau et al. 2004). A comprehensive understanding of the individual contributions of the growing Smad proteome awaits future genetic analysis.

Materials and methods

Derivation of ES cell lines and production of chimeric embryos

N-terminally Flag-tagged human (h) *Smad2(FL)*, *Smad2(Δ30)*, and *Smad3* cDNA cassettes (Yagi et al. 1999) were subcloned into a modified version of pCAGGS (Niwa et al. 1991). Each expression vector was linearized and electroporated into either the KT11 or KT15 *Smad2[Robm1]* homozygous ES cell lines that constitutively express the *Rosa26 lacZ* reporter (Tremblay et al. 2000). Hygromycin-resistant clones were initially screened for expression of the introduced construct by flow cytometry. For cytoplasmic staining, cells were treated with 10% formalin and extensively washed with PBS containing 0.1% saponin. STO feeder cells were eliminated from the analysis by appropriate gating. Primary antibodies (Ab) included mouse anti-Smad2/3 (Transduction), which cross-reacts with hSmad2/3, and, as a control, mouse anti-β-galactosidase (Promega). The secondary reagent used was FITC-conjugated goat anti-mouse IgG (H + L; Caltag). Fluorescence was analyzed using a FACScan flow cytometer, and the data are displayed as cell number versus log fluorescence. The cell lines developed in this study (Table 1) as well as the control ES cell lines R26.1 (Varlet et al. 1997) and KT15 were separately injected into host ICR (Taconic) blastocysts (Nagy et al. 2003). Manipulated embryos were recovered between E9.0 and E10.0, X-gal stained in whole mount, embedded in paraffin wax, and serially sectioned according to standard procedures.

Generation of Smad2 mutant alleles

Deletion of Smad2 coding exon 3. A 488-bp KpnI–NcoI fragment of the *Smad2* gene that contains the alternatively spliced

exon 3 was replaced with a *loxP*-flanked PGK-*neomycin* (*neo*) resistance cassette by homologous recombination in ES cells. The targeting vector contains 2.4-kb 5′ and 1.3-kb 3′ homology arms. Linearized vector was electroporated into CCE ES cells, and drug-resistant colonies were genotyped by Southern blot analysis using either a 5′-external 1.6-kb EcoRI–HindIII probe (Fig. 3B; Nomura and Li 1998) or a 3′-internal 962-bp SpeI–EcoRI probe with KpnI digestion of genomic DNA. Four correctly targeted clones out of 2200 were identified.

Introduction of Smad3 *coding sequences into the mouse* Smad2 *locus.* The ATG-containing exon 2 of *Smad2* was disrupted by homologous recombination with a vector containing 2.7 kb of 5′ homology, followed by an N-terminally Flag-tagged human *Smad3* cDNA (above), a triply repeated SV40 polyadenylation cassette to ensure transcriptional termination (Maxwell et al. 1989), a *loxP*-flanked PGK-*neo* cassette, and 1.86 kb of 3′ homology. A PGK-*dta* cassette was used for negative selection. Linearized vector was electroporated into CCE ES cells, and drug-resistant colonies were genotyped by Southern blot analysis with a 5′-external probe as previously described (Waldrip et al. 1998) or a 3′-internal 587-bp EcoRV–SpeI probe with SpeI digestion of genomic DNA. Five correctly targeted clones out of 600 were identified.

For both *Smad2* alleles, targeted clones were transiently transfected with Cre recombinase. Loss of the *neo* cassette was verified by Southern analysis with a 630-bp PstI–XbaI *neo* probe. At least two independently targeted ES cell clones were used to generate germline chimeric mice.

Mouse strains and genotyping procedures

Smad2$^{\Delta exon3}$/+ and *Smad2*^{hSmad3ki}/+ (129/Sv//Ev × C57BL/6)$_{F1}$ mice were outcrossed to ICR, and then each line was maintained by intercrossing. Outbred *Smad2*$^{\Delta exon3/\Delta exon3}$ and *Smad3*$^{null/null}$ mice were intercrossed. The PCR primers used to detect the *Smad2*$^{\Delta exon3}$ allele were Δ2-1, 5′-GAACTGATCCTCCTGTT TCC-3′; and Δ2-B5, 5′-TGGCACGCTGATA CTTTACACG-3′; with an annealing temperature of 59°C for 35 cycles (Fig. 3C). The primers used to detect the *Smad2*^{hSmad3ki} allele were Smad2com, 5′-AGTTAATTGCCCAGAGCGTTGACA-3′; FLAG1, 5′-CATCGTCCTTGTAGTCCATG-3′; and B5, 5′-CACTTTTC TTCCTGTCCATTCTGC-3′; with an annealing temperature of 58°C for 35 cycles (Fig. 4C). Detection of the *Smad3*^{null} allele was as previously described (Dunn et al. 2004).

Ribonuclease protection assays

Total RNA from CCE ES cells, mouse embryos at various stages, or adult organs was prepared using Trizol (Invitrogen). Ribonuclease protection assays were performed on 10 μg of total RNA (RPA III kit; Ambion). The following probes were used: a 3′ 221-bp BglII–BamHI fragment from mouse *Smad2* (Waldrip et al. 1998; Dunn et al. 2004); a 5′ 325-bp PCR-derived probe from the mouse *Smad2* cDNA that protects coding exons 2 and 3 (5′-primer, 5′-CGGAATTCGC<u>ATG</u>TCGTCCATCTTGCCA TT-3′; 3′-primer, 5′-CGGGATCCGCTGGTTTGTTCAGAGA AGC); a 3′ 237-bp HincII fragment from mouse *Smad3*; and a 3′-probe corresponding to the last 255 bp of human *Smad3*. RPA products were separated on 5% PAGE gels, exposed to film, and quantitated by PhosphorImager.

B- and T-cell stimulation with hTGFβ

Spleen cell suspensions depleted of erythrocytes were incubated at 37°C in B-cell medium containing LPS (50 μg/mL; Difco Laboratories) and hTGFβ (5 ng/mL; R&D Systems) at a density of 5 × 10⁷ cells/mL for 48 h. Thymocytes were incubated at 37°C in T-cell medium containing hTGFβ-1 (5 ng/mL) for 1 h at a density of 4 × 10⁷ cells/mL. Cell suspensions were harvested, washed with PBS containing 2% FCS and antibiotics, and lysed in buffer containing protease and phosphatase inhibitors (Sigma). After incubation on ice, extracts were cleared of nuclei and debris by centrifugation.

Protein purification and Western blotting

Protein extracts (50 μg) derived from ES cells, COS cells transfected with the pCAGGS-N-Flag-hSmad expression constructs (Lipofectamine 2000), and B and T cells were mixed with an equal volume of 2× Laemmli buffer and then separated on a 10% SDS-PAGE gel followed by transfer to nitrocellulose (Protran). ES cell and COS cell extracts were analyzed with the anti-Flag M5 monoclonal (Sigma) in combination with an HRP-conjugated sheep anti-mouse secondary (Amersham); stimulated B, T, and HepG2 cells with a rabbit anti-phospho-smad2 (Cell Signaling), followed by HRP-conjugated donkey anti-rabbit secondary (Amersham). Blots were developed by chemiluminescence using ECL (Amersham).

References

Arai, T., Akiyama, Y., Okabe, S., Ando, M., Endo, M., and Yuasa, Y. 1998. Genomic structure of the human Smad3 gene and its infrequent alterations in colorectal cancers. *Cancer Lett.* **122:** 157–163.

Beck, S., Le Good, J.A., Guzman, M., Ben Haim, N., Roy, K., Beermann, F., and Constam, D.B. 2002. Extraembryonic proteases regulate Nodal signalling during gastrulation. *Nat. Cell Biol.* **4:** 981–985.

Bourillot, P.Y., Garrett, N., and Gurdon, J.B. 2002. A changing morphogen gradient is interpreted by continuous transduction flow. *Development* **129:** 2167–2180.

Brennan, J., Lu, C.C., Norris, D.P., Rodriguez, T.A., Beddington, R.S., and Robertson, E.J. 2001. Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411:** 965–969.

Cheng, S.K., Olale, F., Bennett, J.T., Brivanlou, A.H., and Schier, A.F. 2003. EGF-CFC proteins are essential coreceptors for the TGF-β signals Vg1 and GDF1. *Genes & Dev.* **17:** 31–36.

Chou, W.C., Prokova, V., Shiraishi, K., Valcourt, U., Moustakas, A., Hadzopoulou-Cladaras, M., Zannis, V.I., and Kardassis, D. 2003. Mechanism of a transcriptional cross talk between transforming growth factor-β-regulated Smad3 and Smad4 proteins and orphan nuclear receptor hepatocyte nuclear factor-4. *Mol. Biol. Cell* **14:** 1279–1294.

Datto, M.B., Frederick, J.P., Pan, L., Borton, A.J., Zhuang, Y., and Wang, X.F. 1999. Targeted disruption of Smad3 reveals an essential role in transforming growth factor β-mediated signal transduction. *Mol. Cell. Biol.* **19:** 2495–2504.

Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M.,

Goodstein, D.M., et al. 2002. The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. *Science* **298:** 2157–2167.

Dennler, S., Huet, S., and Gauthier, J.M. 1999. A short amino-acid sequence in MH1 domain is responsible for functional differences between Smad2 and Smad3. *Oncogene* **18:** 1643–1648.

Derynck, R., Akhurst, R.J., and Balmain, A. 2001. TGF-β signaling in tumor suppression and cancer progression. *Nat. Genet.* **29:** 117–129.

Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F., and Wrana, J.L. 2003. Distinct endocytic pathways regulate TGF-β receptor signalling and turnover. *Nat. Cell Biol.* **5:** 410–421.

Duboc, V., Rottinger, E., Besnardeau, L., and Lepage, T. 2004. Nodal and BMP2/4 signaling organizes the oral–aboral axis of the sea urchin embryo. *Dev. Cell* **6:** 397–410.

Dunn, N.R., Vincent, S.D., Oxburgh, L., Robertson, E.J., and Bikoff, E.K. 2004. Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development* **131:** 1717–1728.

Faure, S., Lee, M.A., Keller, T., ten Dijke, P., and Whitman, M. 2000. Endogenous patterns of TGFβ superfamily signaling during early *Xenopus* development. *Development* **127:** 2917–2931.

Feng, X.H., Lin, X., and Derynck, R. 2000. Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-β. *EMBO J.* **19:** 5178–5193.

Han, S.U., Kim, H.T., Seong, D.H., Kim, Y.S., Park, Y.S., Bang, Y.J., Yang, H.K., and Kim, S.J. 2004. Loss of the Smad3 expression increases susceptibility to tumorigenicity in human gastric cancer. *Oncogene* **23:** 1333–1341.

Hata, A., Lo, R.S., Wotton, D., Lagna, G., and Massagué, J. 1997. Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4. *Nature* **388:** 82–87.

Heyer, J., Escalante-Alcalde, D., Lia, M., Boettinger, E., Edelmann, W., Stewart, C.L., and Kucherlapati, R. 1999. Postgastrulation Smad2-deficient embryos show defects in embryo turning and anterior morphogenesis. *Proc. Natl. Acad. Sci.* **96:** 12595–12600.

Inman, G.J., Nicolas, F.J., and Hill, C.S. 2002. Nucleocytoplasmic shuttling of Smads 2, 3, and 4 permits sensing of TGF-β receptor activity. *Mol. Cell* **10:** 283–294.

Jayaraman, L. and Massagué, J. 2000. Distinct oligomeric states of SMAD proteins in the transforming growth factor-β pathway. *J. Biol. Chem.* **275:** 40710–40717.

Kasahara, M., Suzuki, T., and Pasquier, L.D. 2004. On the origins of the adaptive immune system: Novel insights from invertebrates and cold-blooded vertebrates. *Trends Immunol.* **25:** 105–111.

Kjellman, C., Honeth, G., Jarnum, S., Lindvall, M., Darabi, A., Nilsson, I., Edvardsen, K., Salford, L.G., and Widegren, B. 2004. Identification and characterization of a human smad3 splicing variant lacking part of the linker region. *Gene* **327:** 141–152.

Kretschmer, A., Moepert, K., Dames, S., Sternberger, M., Kaufmann, J., and Klippel, A. 2003. Differential regulation of TGF-β signaling through Smad2, Smad3 and Smad4. *Oncogene* **22:** 6748–6763.

Kurisaki, A., Kose, S., Yoneda, Y., Heldin, C.H., and Moustakas, A. 2001. Transforming growth factor-β induces nuclear import of Smad3 in an importin-β1 and Ran-dependent manner. *Mol. Biol. Cell* **12:** 1079–1091.

Labbé, E., Silvestri, C., Hoodless, P.A., Wrana, J.L., and Attisano, L. 1998. Smad2 and Smad3 positively and negatively regulate TGF β-dependent transcription through the forkhead DNA-binding protein FAST2. *Mol. Cell* **2:** 109–120.

Lareau, L.F., Green, R.E., Bhatnagar, R.S., and Brenner, S.E. 2004. The evolving roles of alternative splicing. *Curr. Opin. Struct. Biol.* **14:** 273–282.

Lebrun, J.J., Takabe, K., Chen, Y., and Vale, W. 1999. Roles of pathway-specific and inhibitory Smads in activin receptor signaling. *Mol. Endocrinol.* **13:** 15–23.

Liu, F. 2003. Receptor-regulated Smads in TGF-β signaling. *Front Biosci.* **8:** s1280–s1303.

Liu, Y., Festing, M., Thompson, J.C., Hester, M., Rankin, S., El-Hodiri, H.M., Zorn, A.M., and Weinstein, M. 2004. Smad2 and Smad3 cooperatively regulate craniofacial and endodermal development. *Dev. Biol.* **270:** 411–426.

Massagué, J. and Wotton, D. 2000. Transcriptional control by the TGF-β/Smad signaling system. *EMBO J.* **19:** 1745–1754.

Massagué, J., Blain, S.W., and Lo, R.S. 2000. TGFβ signaling in growth control, cancer, and heritable disorders. *Cell* **103:** 295–309.

Matsuura, I., Denissova, N.G., Wang, G., He, D., Long, J., and Liu, F. 2004. Cyclin-dependent kinases regulate the antiproliferative function of Smads. *Nature* **430:** 226–231.

Maxwell, I.H., Harrison, G.S., Wood, W.M., and Maxwell, F. 1989. A DNA cassette containing a trimerized SV40 polyadenylation signal which efficiently blocks spurious plasmid-initiated transcription. *Biotechniques* **7:** 276–280.

Miyazawa, K., Shinozaki, M., Hara, T., Furuya, T., and Miyazono, K. 2002. Two major Smad pathways in TGF-β superfamily signalling. *Genes Cells* **7:** 1191–1204.

Morokuma, J., Ueno, M., Kawanishi, H., Saiga, H., and Nishida, H. 2002. HrNodal, the ascidian nodal-related gene, is expressed in the left side of the epidermis, and lies upstream of HrPitx. *Dev. Genes Evol.* **212:** 439–446.

Nagy, A., Gertsentein, M., Vintersten, K., and Behringer, R. 2003. *Manipulating the mouse embryo.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Newfeld, S.J., Wisotzkey, R.G., and Kumar, S. 1999. Molecular evolution of a developmental pathway: Phylogenetic analyses of transforming growth factor-β family ligands, receptors and Smad signal transducers. *Genetics* **152:** 783–795.

Nishita, M., Ueno, N., and Shibuya, H. 1999. Smad8B, a Smad8 splice variant lacking the SSXS site that inhibits Smad8-mediated signalling. *Genes Cells* **4:** 583–591.

Niwa, H., Yamamura, K., and Miyazaki, J. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108:** 193–199.

Nomura, M. and Li, E. 1998. Smad2 role in mesoderm formation, left–right patterning and craniofacial development. *Nature* **393:** 786–790.

Norris, D.P., Brennan, J., Bikoff, E.K., and Robertson, E.J. 2002. The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. *Development* **129:** 3455–3468.

Osada, S.I., Saijoh, Y., Frisch, A., Yeo, C.Y., Adachi, H., Watanabe, M., Whitman, M., Hamada, H., and Wright, C.V. 2000. Activin/Nodal responsiveness and asymmetric expression of a *Xenopus* Nodal-related gene converge on a FAST-regulated module in intron 1. *Development* **127:** 2503–2514.

Pardali, K., Kurisaki, A., Moren, A., ten Dijke, P., Kardassis, D., and Moustakas, A. 2000. Role of Smad proteins and transcription factor Sp1 in p21[Waf1/Cip1] regulation by transforming growth factor-β. *J. Biol. Chem.* **275:** 29244–29256.

Parisi, S., D'Andrea, D., Lago, C.T., Adamson, E.D., Persico, M.G., and Minchiotti, G. 2003. Nodal-dependent Cripto signaling promotes cardiomyogenesis and redirects the neural fate of stem cells. *J. Cell Biol.* **163:** 303–314.

Piek, E., Ju, W.J., Heyer, J., Escalante-Alcalde, D., Stewart, C.L., Weinstein, M., Deng, C., Kucherlapati, R., Bottinger, E.P.,

and Roberts, A.B. 2001. Functional characterization of transforming growth factor β signaling in Smad2- and Smad3-deficient fibroblasts. *J. Biol. Chem.* **276:** 19945–19953.

Pierreux, C.E., Nicolas, F.J., and Hill, C.S. 2000. Transforming growth factor β-independent shuttling of Smad4 between the cytoplasm and nucleus. *Mol. Cell. Biol.* **20:** 9041–9054.

Reguly, T. and Wrana, J.L. 2003. In or out? The dynamics of Smad nucleocytoplasmic shuttling. *Trends Cell Biol.* **13:** 216–220.

Saijoh, Y., Adachi, H., Sakuma, R., Yeo, C.Y., Yashiro, K., Watanabe, M., Hashiguchi, H., Mochida, K., Ohishi, S., Kawabata, M., et al. 2000. Left–right asymmetric expression of lefty2 and nodal is induced by a signaling pathway that includes the transcription factor FAST2. *Mol. Cell* **5:** 35–47.

Seoane, J., Le, H.V., Shen, L., Anderson, S.A., and Massagué, J. 2004. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* **117:** 211–223.

Shi, Y. and Massagué, J. 2003. Mechanisms of TGF-β signaling from cell membrane to the nucleus. *Cell* **113:** 685–700.

Shi, Y., Wang, Y.F., Jayaraman, L., Yang, H., Massagué, J., and Pavletich, N.P. 1998. Crystal structure of a Smad MH1 domain bound to DNA: Insights on DNA binding in TGF-β signaling. *Cell* **94:** 585–594.

Shimizu, K. and Gurdon, J.B. 1999. A quantitative analysis of signal transduction from activin receptor to nucleus and its relevance to morphogen gradient interpretation. *Proc. Natl. Acad. Sci.* **96:** 6791–6796.

Shiratori, H., Sakuma, R., Watanabe, M., Hashiguchi, H., Mochida, K., Sakai, Y., Nishino, J., Saijoh, Y., Whitman, M., and Hamada, H. 2001. Two-step regulation of left–right asymmetric expression of Pitx2: Initiation by nodal signaling and maintenance by Nkx2. *Mol. Cell* **7:** 137–149.

Sibilia, M., Wagner, B., Hoebertz, A., Elliott, C., Marino, S., Jochum, W., and Wagner, E.F. 2003. Mice humanised for the EGF receptor display hypomorphic phenotypes in skin, bone and heart. *Development* **130:** 4515–4525.

Siegel, P.M. and Massagué, J. 2003. Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer. *Nat. Rev. Cancer* **3:** 807–821.

Song, J., Oh, S.P., Schrewe, H., Nomura, M., Lei, H., Okano, M., Gridley, T., and Li, E. 1999. The Type II Activin receptors are essential for egg cylinder growth, gastrulation, and rostral head development in mice. *Dev. Biol.* **213:** 157–169.

Takenoshita, S., Mogi, A., Nagashima, M., Yang, K., Yagi, K., Hanyu, A., Nagamachi, Y., Miyazono, K., and Hagiwara, K. 1998. Characterization of the MADH2/Smad2 gene, a human Mad homolog responsible for the transforming growth factor-β and activin signal transduction pathway. *Genomics* **48:** 1–11.

Tremblay, K.D., Hoodless, P.A., Bikoff, E.K., and Robertson, E.J. 2000. Formation of the definitive endoderm in mouse is a Smad2-dependent process. *Development* **127:** 3079–3090.

Tsukazaki, T., Chiang, T.A., Davison, A.F., Attisano, L., and Wrana, J.L. 1998. SARA, a FYVE domain protein that recruits Smad2 to the TGFβ receptor. *Cell* **95:** 779–791.

Varlet, I., Collignon, J., and Robertson, E.J. 1997. Nodal expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. *Development* **124:** 1033–1044.

Venter, J., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. 2001. The sequence of the human genome. *Science* **291:** 1304–1351.

Vincent, S.D., Dunn, N.R., Hayashi, S., Norris, D.P., and Rob-

ertson, E.J. 2003. Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes & Dev.* **17:** 1646–1662.

Waldrip, W.R., Bikoff, E.K., Hoodless, P.A., Wrana, J.L., and Robertson, E.J. 1998. Smad2 signaling in extraembryonic tissues determines anterior–posterior polarity of the early mouse embryo. *Cell* **92:** 797–808.

Wall, N.A., Craig, E.J., Labosky, P.A., and Kessler, D.S. 2000. Mesendoderm induction and reversal of left–right pattern by mouse Gdf1, a Vg1-related gene. *Dev. Biol.* **227:** 495–509.

Whitman, M. 1998. Smads and early developmental signaling by the TGFβ superfamily. *Genes & Dev.* **12:** 2445–2462.

Xu, L. and Massagué, J. 2004. Nucleocytoplasmic shuttling of signal transducers. *Nat. Rev. Mol. Cell. Biol.* **5:** 209–219.

Yagi, K., Goto, D., Hamamoto, T., Takenoshita, S., Kato, M., and Miyazono, K. 1999. Alternatively spliced variant of Smad2 lacking exon 3. Comparison with wild-type Smad2 and Smad3. *J. Biol. Chem.* **274:** 703–709.

Yang, X., Letterio, J.J., Lechleider, R.J., Chen, L., Hayman, R., Gu, H., Roberts, A.B., and Deng, C. 1999. Targeted disruption of SMAD3 results in impaired mucosal immunity and diminished T cell responsiveness to TGF-β. *EMBO J.* **18:** 1280–1291.

Yang, X., Chen, L., Xu, X., Li, C., Huang, C., and Deng, C.X. 2001. TGF-β/Smad3 signals repress chondrocyte hypertrophic differentiation and are required for maintaining articular cartilage. *J. Cell Biol.* **153:** 35–46.

Yang, Y.C., Piek, E., Zavadil, J., Liang, D., Xie, D., Heyer, J., Pavlidis, P., Kucherlapati, R., Roberts, A.B., and Bottinger, E.P. 2003. Hierarchical model of gene regulation by transforming growth factor β. *Proc. Natl. Acad. Sci.* **100:** 10269–10274.

Yu, J.K., Holland, L.Z., and Holland, N.D. 2002. An amphioxus nodal gene (AmphiNodal) with early symmetrical expression in the organizer and mesoderm and later asymmetrical expression associated with left–right axis formation. *Evol. Dev.* **4:** 418–425.

Zhu, Y., Richardson, J.A., Parada, L.F., and Graff, J.M. 1998. Smad3 mutant mice develop metastatic colorectal cancer. *Cell* **94:** 703–714.