Villin, an actin bundling protein found in the apical brush border of absorptive tissues, is one of the first structural genes to be transcriptionally activated in the embryonic intestinal endoderm. In the adult, villin is broadly expressed in every cell of the intestinal epithelium on both the vertical axis (crypt to villus tip) and the horizontal axis (duodenum through colon) of the intestine. Here, we document that a 12.4-kilobase region of the mouse villin gene drives high level expression of two different reporter genes (LaCZ and Cre recombinase) within the entire intestinal epithelium of transgenic mice. Deletion of a portion of this transgene results in reduction of β-galactosidase activity in restricted domains of the small intestine (duodenum) and large intestine (cecum). In addition, expression is reduced in the crypt compartment throughout the intestine. Thus, the global expression pattern of villin in the intestine is apparently the consequence of an amalgam of distinct and individual domain-specific control processes. That is, expression of villin in the duodenum and cecum requires different regulatory sequences than the rest of the intestine, and the expression of villin in crypts is regulated by different circuitry than expression of villin on villus tips.

Much has been learned about the process of organogenesis through analysis of the regulation of structural genes that are expressed tissue-specifically. For example, the study of globin gene expression provided important insights into red cell development (1, 2), whereas the study of albumin expression led to an improved understanding of the factors that regulate development of the liver (3). Similarly, we propose that the actin bundling protein, villin, is an excellent model gene for the study of intestinal epithelial organogenesis.

The villin gene is first activated at 9.0 days post coitum (dpc) in the presumptive intestinal hindgut endoderm during gut tube closure (4–6). The early (10 dpc) expression domain of villin rapidly extends throughout the small and large intestinal endoderm and includes the distal stomach (4, 6). At 14–15 dpc, when the intestinal epithelium is dramatically remodeled and villi are formed, villin gene expression is up-regulated (6, 7). At 16 dpc, a sharp, one-cell-thick border of villin expression becomes apparent in the pyloric epithelium; intestinal cells express villin at high levels whereas neighboring stomach cells exhibit low or barely detectable expression of villin (4). Once intestinal crypts are established, villin is expressed in all cells of the crypt-villus axis, with an increasing gradient from crypt to tip (6). Thus, the villin gene appears to respond to the various morphogenetic cues that dictate key events of intestinal development at multiple time points both in utero and postnatally. It is therefore likely that the molecular factors that dictate this pattern of villin expression are the same factors that coordinate these key events of intestinal development.

Villin regulatory sequences have been investigated in two previous studies (8, 9). However, because the in vivo transcriptional start site of villin was initially erroneously mapped to the same exon as the translational start codon (9), an early cell culture study analyzed a villin “promoter” fragment that actually consisted entirely of intronic sequences (9). Later work showed that transcription actually begins at a 5′ non-translated exon that is separated from exon 1 (the ATG-containing exon) by 5.6 kb in the mouse (8, 10). A 9-kb fragment, containing this entire intron plus 3.4 kb of additional upstream 5′-flanking sequence, was recently used to generate transgenic mice (8). This 9-kb fragment drove reporter gene expression in the intestine of adult mice, although four of five founders showed “heterogeneous” expression. Because this analysis was limited to the study of founders, embryonic expression of the transgene was not investigated and the possible effects of transgene mosaicism versus position effects were not assessed (8).

Here, we utilized a larger transgene construct and analyzed both fetal and adult transgene expression in multiple established lines. We report that a 12.4-kb region from the mouse villin gene directs intestine-specific expression of the bacterial β-galactosidase gene in a manner that largely recapitulates the normal expression pattern of villin in the intestine. In 75% of the established lines, including two single copy lines, this transgene directs continuous (not variegated) expression in the entire intestinal epithelium. This suggests that this 12.4-kb transgene contains a locus control region (LCR), a chromatin-
remodeling cis element that facilitates position-independent expression (11–15). We also demonstrate that this 12.4-kb regulatory fragment can be used to direct expression of Cre recombinase in the large and small intestines of transgenic mice.

Analysis of a deletion derivative of this 12.4-kb regulatory fragment revealed the presence of one or more multifunctional enhancers within a 4.2-kb region of the first intron. This enhancer region appears to be at least partially responsible for mediating LCR-like activity and is necessary for high level expression of β-galactosidase in specific regions of the intestine, including the duodenum and Cecum. In addition, the same regulatory fragment can be used to direct expression of Cre recombinase in the cecum. We also demonstrate that this 12.4-kb regulatory fragment is necessary for high level expression of Cre recombinase, which is necessary for high level expression of Cre recombinase in specific regions of the intestine, including the duodenum and Cecum.

EXPERIMENTAL PROCEDURES

Plasmid and Transgene Construction—The 12.4-kb promoter fragment from the villin gene was isolated from cosmid clones kindly provided by Dr. Philippe Gros, McGill University, Montreal, Quebec, Canada. After digestion with EcoRI (6.7 kb 5′ of the transcriptional start site) and XmnI (in exon 1), a 12.4-kb promoter fragment was isolated and ligated into EcoRI and XmnI sites of the pBluescript II SK+ (pBSII SK+) plasmid (Strategene) to create the plasmid pBSII-12.4kbVillin. A deletion in intron 1 of this fragment was generated by digestion of the pBSII-12.4kbVillin plasmid with SfI followed by re-ligation of the plasmid to create the plasmid pBSII-ΔNheVillin. The LacZ coding region was removed from the pCH110 vector (Clontech) by digestion with KpnI and EcoRI and ligated into a modified pGEM7 vector (Promega) in which the XhoI site was replaced with an XmnI site. The LacZ insert was removed from this vector by digestion with XmnI and ligated into XmnI sites in pBSII-12.4kbVillin and pBSII-ΔNheVillin plasmids. This created the pBSII-12.4kbVillinLacZ and pBSII-ΔNheVillinLacZ plasmids. In both final plasmids, the β-galactosidase protein (with the first 39 amino acids removed) is fused to the first 18 amino acids of the villin protein.

To make the pBSII-12.4kbVillin plasmid a more useful tool for expression of any suitable cDNA (including Cre recombinase), the villin ATG and first 18 amino acids were removed. A 2.0-kb region of the villin intron was amplified using a forward primer specific for sequences flanking a unique BamHI site and a reverse primer specific for the first 11 bp of exon 1. The reverse primer in exon 1 excluded the villin ATG and contained a linker with restriction sites for SmaI, EcoRI, EcoRV, and NotI. The fragment was then cloned into the pBSII-12.4kbVillin plasmid and NotI sites were created by excising the pBSII-12.4kbVillinLacZ plasmids. A Cre recombinase cDNA with a metallothionein poly(A) signal was isolated as a XhoI/BamHI fragment from the pBS185 plasmid (Invitrogen) and cloned into pBS (Invitrogen) to create the pBS-Cre plasmid. A SalI/NotI fragment from the pBSII-12.4kbVillin/ΔATG plasmid was cloned into the SalI/NotI sites of the pBS-Cre plasmid to create the pBS-12.4kbVillinCre plasmid. All engineered regions were sequenced to ensure fidelity.

Generation of Transgenic Mice—DNA fragments containing villin regulatory sequences linked to LacZ were prepared by digestion of pBSII-12.4kbVillinLacZ and pBSII-ΔNheVillinLacZ plasmids with EcoRI and XhoI. The 12.4kbVillinCre transgene construct was liberated from the pBS-VillinCre vector sequences using XhoI, and contained a linker with restriction sites for KpnI and SmaI. The reverse primer in exon 1 excluded the villin ATG for 0.01% Triton X-100 in PBS at 4 °C with agitation, washed with ice-cold PBS, then homogenized on ice in 6 volumes of homogenization buffer (PBS containing 40 μM phenylmethylsulfonlfuoride, 3.13 mg/ml benzamidine, 20 μg/ml leupeptin, 20 μg/ml E64) per milligram of tissue. Protein concentrations were determined by triplicate in the Lowry method using the Bio-Rad DC Protein Assay kit. β-Galactosidase assays were performed in triplicate with 10–75 μg of total protein tissue extracts by diluting with reaction buffer (0.1 M sodium phosphate buffer, pH 7.4, 1 mM MgCl2) containing 0.88 mg/ml O-nitrophenyl-β-D-galactopyranoside) to a total volume of 300 μl. Samples were incubated at 37 °C for 30 or 50 min (for kidney and yolk sac extracts), and reactions were stopped by addition of 0.5 ml of 1 N Na2CO3. Samples were then read on a spectrophotometer for absorbance at 420 nm.

Southern Blot Analysis of Transgene Copy Number—Genomic DNA (6–8 μg) isolated from a tail tip liver was digested with SfI and XhoI restriction enzymes and subjected to electrophoresis on a 1.25% agarose gel. Gels were blotted overnight to Hybond N+ membranes (Amersham Biosciences). Membranes were cross-linked by exposing to 1200 μJ of UV radiation in a UV Stratalinker (Stratagene) and prehybridized for at least 1 h at 60 °C with 10 ml (per 200 cm2 of membrane) of ExpressHyb solution (Clontech) supplemented with 150 μg/ml single-stranded salmon sperm DNA and with 0.2% w/v SDS. Hybridization was performed by PCR amplification of a region in intron 1 (between 5′-CAATGGAGGGTTCTTTTTGTG-3′ and 5′-AACATGGCTTCTATGGGTC-3′). Labeled probe was denatured at 95 °C for 5 min, quick-cooled on ice, diluted to 2 × 106 cpm/ml with ExpressHyb solution (Clontech) supplemented with 150 μg/ml single-stranded salmon sperm, and added to membranes (10 ml per 200 cm2 of membrane). Hybridization was carried out for 14–16 h at 60 °C. Band intensity was assessed on a PhosphorImager (Amersham Biosciences).

RT-PCR Analysis of Transgene mRNA Expression—Total RNA was isolated from tissues using the TRIzol reagent (Invitrogen), and concentrations were determined using a spectrophotometer. One microgram of total RNA was used for reverse transcriptase (RT)-mediated cDNA synthesis using random hexamer primers. To assess the RT reaction and quality of synthesized cDNA, PCR amplification of the hypoxanthine phosphoribosyl transferase cDNA was performed using the specific primers: forward 5′-CAAGAGCTAGAAGACCTGCG-3′ and reverse 5′-GCTTGTGAAAAAGGACCTC-3′. Primers for the transgene mRNA were designed to flank the villin transgenic intron 1 with the forward primer located in intron 1 (5′-TCACCGGAGGCAAACCTACCCAAG-3′) and the backward primer located in the coding region of LacZ (5′-TCACTCCAAACGCGCCATATCCCC-3′). These primers amplify the same 744-bp fragment from both transgenes. PCR amplification was performed in a thermocycler for 35 cycles with 30-s cycles of denaturation at 94 °C and 45 s of polymerase extension at 72 °C. Annealing temperatures were performed as follows: for the first five cycles the annealing step was programmed at 68 °C for 30 s followed by 10 cycles.
in the villin gene. A, schematic of the endogenous villin locus. The locus contains an upstream untranslated exon (UTR) and a 5.6-kb region corresponding to intron 1. The start codon for protein translation is located in exon 1. The constructs used in transient transfection assays. C, relative luciferase activity for each construct in CaCo2 intestinal cells and Rat2 fibroblasts. Luciferase activity was assayed for each construct, and the results were normalized to β-galactosidase activity. Transfections were repeated four to six times for each construct. The error bars represent standard error of the mean.

### RESULTS

**Transient Transfection of Mouse Villin Gene Sequences**—As shown in Fig. 1A, the villin gene locus contains a 5'-untranslated region (UTR) that is separated from the first coding exon by a 5.6-kb intron. The constructs used in these transfection experiments are shown in Fig. 1B, whereas luciferase assay results (normalized by co-transfected β-galactosidase activity) are shown in Fig. 1C.

A construct containing 1236 bp of 5'-flanking sequence (−1 to −1236 from the transcriptional start site) reliably directs luciferase expression at high levels in CaCo2 cells, and at 5- to 6-fold lower levels in Rat2 fibroblasts (−1236, Fig. 1C). Removal of the 5'-most 682 bp significantly reduces transcriptional activity in both cell types (−554, Fig. 1, B and C). Removal of an additional 88 bp (−466) results in a further 2-fold reduction of reporter activity in CaCo2 cells. Thus, critical activating sequences appear to lie between −1236 and −554 bp, as well as between −554 and −466 bp. This conforms and refines earlier studies indicating the presence of activators between −3.5 kb and −480 bp and between −480 and −100bp of the mouse villin gene (8).

Addition of the 5.6-kb first intron to the −554 construct enhances promoter activity nearly 4-fold in CaCo2 cells, whereas promoter activity is unchanged in Rat2 fibroblasts (compare −544 to −554/intron, Fig. 1, B and C). Interestingly, Pinto et al. (8) also tested the effect of adding this intron and noted a 2-fold enhancement when the intron was linked to 3.5 kb of upstream sequence but no effect (or slight diminution) when the intron was linked to −480 bp of upstream sequence. Together, the data from the two studies indicate that elements within the first intron may interact with or synergize with 5'-flanking sequences between −480 and −544 bp.

Deletion of a 4.2-kb region between two NheI sites in intron 1 results in a 2-fold reduction in reporter activity in CaCo2 cells, compared with the construct containing the entire intron (−554/Nhe, Fig. 1, B and C). In contrast, this deletion increases expression in Rat2 fibroblasts by more than 4-fold. Thus, enhancer sequences within this 4.2-kb region appear to facilitate expression in intestinal-like cells while repressing expression in non-intestinal cells.

To further explore the in vivo function of the putative intestinal cell-type-specific enhancer in intron 1 of the villin gene, two transgene constructs were prepared in which villin regulatory sequences were linked to a β-galactosidase reporter. Because earlier work (8) had shown that a transgene construct encompassing 9 kb of the mouse villin gene, including the first intron, was not sufficient to impart position-independent expression, we began with a larger segment of 5' sequence. A transgene construct, 12.4KbVilLacZ, was prepared containing 6.7 kb of sequence upstream of the transcriptional start site, the untranslated exon (exon −1), the entire first intron, and the first 65 base pairs of exon 1, fused in-frame to the bacterial β-galactosidase gene (Fig. 2A). A second transgene construct, ΔNheVilLacZ, was also generated, in which the same 5'-flanking sequences were included, but the 4.2-kb region of intron 1 containing the putative intronic enhancer identified in cell culture studies was deleted (Fig. 2A).

β-Galactosidase Expression in Transgenic Mice—Eight founders carrying the 12.4KbVilLacZ transgene and five founders carrying the ΔNheVilLacZ transgene were bred to obtain transgenic lines for analysis. At 4 to 8 weeks of age, tissues from F1 mice obtained from founders crossed to C57BL/6 mice were analyzed for β-galactosidase expression by staining with X-gal. Tissues analyzed included small intestine, colon, kidney, heart, lung, liver, spleen, skeletal muscle, stomach, thymus, thyroid, brain, uterus, ovaries, and testes.

In all eight lines of mice carrying the 12.4KbVilLacZ construct, β-galactosidase activity was detectable exclusively in the small intestine and colon. In most lines, the highest intensity of β-galactosidase staining was seen in the duodenum with progressively diminished staining in the ileum and colon (Fig. 2B), a gradient pattern that was confirmed by a quantitative β-galactosidase enzymatic solution assay (data not shown). In one very highly expressing line (Fig. 2C), β-galactosidase activity was similar along the entire horizontal axis of the intestine. The decreasing gradient of villin expression seen in most lines mimics the expression pattern of villin in the small intestine and colon seen in genetically modified mice in which a β-galactosidase cDNA was inserted into the endogenous villin locus (10). Thus, the 12.4-kb regulatory fragment contains the necessary cis sequences for the cephalocaudal expression pattern of villin in the intestine. The pyloric border of β-galactosidase expression was also appropriately abrupt. No expression of the reporter was observed in stomach, but adjacent intestinal cells were intensely stained (Fig. 2, G and H).

In six of eight lines, including two lines that contained a single copy of the transgene (see below), β-galactosidase ex-
expression was continuous throughout the small and large intestine, showing no evidence of mottled or variegated expression (Fig. 2, B and C). In the intestine, position effect variegation is expected to cause mottled expression in the adult, because crypts are monoclonal (18). Stochastic silencing of the transgene in some crypts and activation in others would result in an all or none expression pattern with some crypts being totally β-galactosidase-positive and others β-galactosidase-negative. This type of patchy expression pattern was observed in two of the eight lines studied; both were high copy number lines (see below). The fact that only 25% of 12.4KbVilLacZ lines showed evidence of patchy expression while 80% of founders carrying a transgene driven by a 9-kb villin fragment (8) exhibited this pattern suggests that the extra 3.4 kb of 5′-flanking sequence used here are important for suppression of chromosomal position effects.

The expression of β-galactosidase was also analyzed in 4- to 8-week-old ΔNheVilLacZ transgenic mice. Highly variegated expression was seen in the intestine in two of five lines (N3, N5), whereas two lines (N1, N4) showed no expression at all in any tissue, including the intestine. One line (N2), exhibited non-variegated expression, except in the duodenum just distal to the stomach (Fig. 2D, asterisk), where expression was both reduced and variegated. The lack of expression in some ΔNheVilLacZ lines and variegated expression in others suggests that this transgene is susceptible to position effect silencing. This is in dramatic contrast to the 12.4KbVilLacZ transgenic mice in which variegated expression was only observed in two out of eight lines. Thus, the deleted 4.2-kb region of intron 1 may contain sequences important for LCR-like activity.

Strikingly, in all three ΔNheVilLacZ lines that expressed β-galactosidase, expression was most severely reduced in the proximal duodenum (Fig. 2, D–F, asterisks). This indicates that, in addition to possible LCR-like function, the 4.2-kb region of intron 1 is also required for region-specific expression in the duodenum. Close examination of the pyloric region in 12.4KbVilLacZ (Fig. 2, G and H) and in ΔNheVilLacZ mice (Fig. 2I) shows that sequences important for the establishment of the discrete epithelial border between stomach and intestine are present in both constructs. Expression was not detected in the stomach in any of the ΔNheVilLacZ lines or the 12.4KbVilLacZ lines.

**β-Galactosidase Expression during Development of Transgenic Embryos—β-Galactosidase activity was evaluated at 10.5, 12.5, and 14.5 dpc in multiple F2 embryos representing three separate lines for each transgene. In both groups of transgenic mice, expression was first detectable in the embryonic midgut and hindgut at 12.5 dpc (data not shown). This is 3 days later than that observed in the LacZ knock-in model (4) and indicates that some critical sequences necessary for transgene activation at the proper time are missing from both constructs. For 12.4KbVilLacZ transgenic mice, embryonic expression was restricted to the midgut and hindgut with expression decreasing distally along the axis of the hindgut (Fig. 3A), identical to the embryonic expression pattern seen in the Villin LacZ knock-in model (4). Interestingly, the further developmentally programmed postnatal up-regulation of β-galactosidase (concomitant with endogenous villin up-regulation) apparently causes this gradient to disappear in adults of this particular line; in other transgenic lines the gradient is still visible in the adult (Fig. 2, B versus C).

In ΔNhe embryos, however, expression was consistently diminished in the presumptive proximal duodenal area just distal to the stomach (Fig. 3B, asterisk). Thus, in both embryos and adults carrying the ΔNhe transgene, β-galactosidase expression is reduced in the duodenum. Also apparent in Fig. 3B is the loss of β-galactosidase staining in the presumptive cecum of ΔNhe mice. Like the reduction in duodenal staining, this is a reproducible finding, seen in three independent ΔNhe lines. However, unlike the duodenal pattern, where reduced staining persisted into adulthood, cecal expression was partially regained in adults (data not shown).

Although the endogenous villin gene is also expressed at high levels in the yolk sac and in the proximal tubules of the kidney (4, 6, 7), neither the 12.4KbVilLacZ nor the ΔNheVilLacZ transgene was expressed in these tissues. This was determined by β-galactosidase solution assay and by RT-PCR analysis (data not shown). Thus, the 12.4-kb regulatory frag-
ment appears to lack cis elements necessary for both kidney and yolk sac expression. Furthermore, no β-galactosidase activity was detected (above background staining) in any other developing tissue of 12.4KbVilLacZ embryos at any of these stages. F2 embryos from ΔNheVilLacZ transgenic mice, however, demonstrated ectopic reporter expression in tissues such as the developing brain and limbs (data not shown). This is in accord with the transient transfection data indicating that sequences within the 4.2-kb Neh1 fragment are required for suppression of expression in Rat2 fibroblasts.

Copy Number-dependent Expression of the 12.4KbVilLacZ Transgene—Judging from the continuous expression pattern and lack of variegation in the majority (75%) of the 12.4KbVilLacZ founder lines and the contrasting highly variegated expression seen in most ΔNhe lines, we predicted that the mouse villin gene contains an LCR-like activity located in intron 1. To test this prediction, we examined copy number-dependent expression, a hallmark of LCR activity. β-Galactosidase expression levels were determined by solution assay in extracts of whole intestine from each line. Genomic DNA from the six non-variegated 12.4KbVilLacZ lines was loaded onto an agarose gel, left to right, in the rank order of β-galactosidase activity (Fig. 4B). DNA from the two variegated lines (12.4-6 and 12.4-7), both of which expressed β-galactosidase at lower levels than the continuous lines, was loaded in the last two lanes (marked PEV). Copy numbers were determined by scanning the Southern blot of this gel (Fig. 4B, left) on a PhosphorImager, and results were plotted against the level of β-galactosidase activity (left to right, lowest to highest). DNA from the two variegated 12.4KbVilLacZ lines, which express β-galactosidase at lower levels than the continuously staining lines, was loaded in the last two lanes (marked PEV). Copy number (reported below each lane) was determined by densitometric analysis of the blots after normalizing the band from the endogenous villin locus to two copies for each lane. Lines 12.4-5 and 12.4-3 were single copy lines. Two of five ΔNheVilLacZ lines (N1 and N4) had no detectable β-galactosidase expression (NE). Two lines (N3 and N5) exhibited mottled or patchy expression throughout the intestine, indicative of position effect variegation (PEV); these two lines had similar β-galactosidase activity. Finally, one line, N2, exhibited continuous expression at a high level, although expression was variegated in the duodenum. C, correlation between copy number and β-galactosidase activity in 12.4KbVilLacZ transgenic lines, as measured in tissue extracts in duodenum (duod.), ileum (il.), and colon (col.) (average $r^2 = 0.723$).

Copy number dependence was similarly assessed in ΔNheVilLacZ lines. As shown in the right panel of Fig. 4B, one of the two non-expressing lines (NE) contained a high transgene copy number of 32, whereas the other contained a relatively low copy number of 6. The two variegated lines (PEV) expressed β-galactosidase at similar levels, but contained different copy numbers (5 and 10 copies). Furthermore, two of five ΔNheVilLacZ lines failed to express at all, indicative of 100% silencing in these chromatin environments. Thus, variegated expression or failure to express was observed in the majority of ΔNheVilLacZ lines (4 of 5). These data suggest that the ΔNhe deletion in intron 1 removes sequences that suppress position effect silencing.

Heterogeneous Staining on Villi and Reduced Expression of β-Galactosidase in the Crypt Compartment of ΔNheVilLacZ Mice—To assess whether these transgenes drive expression...
throughout the vertical axis of the intestine, we examined β-galactosidase staining in cryosections of intestine from transgenic mice. In mice carrying the 12.4KbVilLacZ construct, continuous expression of LacZ was observed in both crypt and tip compartments (Fig. 5, A–C) of the small and large intestines, with slightly lower expression in crypts, in concordance with the reported expression gradient of villin along the crypt-villus tip axis (6).

In lines carrying the ΔNheVilLacZ transgene two differences in epithelial staining were observed. First, the pattern of staining on the villi was heterogeneous, with neighboring cells exhibiting quite different levels of β-galactosidase staining (Fig. 5, D–F). This type of expression pattern has been seen earlier for the endogenous L-FABP gene (Fabpl) (19) and for ILBP transgenes (20). A second interesting difference seen only in ΔNhe mice was that β-galactosidase staining was severely reduced in crypts along the entire cephalocaudal axis of the intestine (Fig. 5, D–F). To confirm that reduced crypt staining was not simply due to an overall reduction in staining in the epithelium to levels that leave crypt cells below the limits of detection, we compared staining in N2 (Fig. 5E) and N5 mice (Fig. 5F) with that seen in 12.4KbVilLacZ line 12.4-3, a single copy line. Although the 12.4-3 line exhibits equal or lower intensity of β-galactosidase staining on villi than either N2 or N5, its crypts are still clearly β-galactosidase-positive (Fig. 5C), whereas those of both ΔNhe lines are strikingly negative. Thus, the data indicate that the 4.2-kb NheI fragment contains a crypt-specific enhancer.

To determine when, during crypt development, cells with reduced β-galactosidase are first observed, we examined the expression of β-galactosidase in postnatal day 0 (p0), postnatal day 7 (p7), and adult (>4-week-old) intestines. In newborn mice, before the emergence of crypts, all villus and intervillus cells of the small intestine are strongly β-galactosidase-positive in both ΔNheVilLacZ and 12.4KbVilLacZ newborn mice (Fig. 6, A and B). However, in p7 ΔNhe mice, the first few invaginating crypt cells of the small intestine show reduced or absent β-galactosidase staining (Fig. 6C), whereas invaginating crypts of 12.4KbVilLacZ mice are strongly β-galactosidase-positive (Fig. 6D). In adult ΔNhe mice, a clear boundary is visible between the β-galactosidase-positive villus tip cells and the adjacent crypt compartment where β-galactosidase staining is reduced or absent (Fig. 6E, arrow). In contrast, all crypt cells of the 12.4KbVilLacZ lines are β-galactosidase-positive (Fig. 6F). The same pattern is also seen in the large intestine: β-galactosidase expression in all cells at all stages in 12.4KbVilLacZ mice (Fig. 6, G and H); reduction in expression specifically in the crypt compartment of adult, but not newborn ΔNheVilLacZ mice (Fig. 6, I and J). These data indicate that a different genetic program for the regulation of villin expression is present in cells of the crypts and cells of the villi. The crypt program, but not the villus program, requires cis elements within the 4.2-kb deleted region.

Spatial Conservation of cis Elements in the Proximal Promoter and Intron 1 Region among Mice and Humans—The nucleotide sequence of the human villin gene was obtained from the NCBI data base of high throughput human genomic sequence. Intron 1 of the human villin gene was cloned from total genomic DNA, sequenced, and found to be identical with sequence from the NCBI data base. The proximal promoter region (obtained by PCR of BAC clone RPCI-11-378A13, GenBank™ accession number AC021016, kindly provided by the BACPAC Resource Center at the Children’s Hospital Oakland Research Institute, Oakland, CA) of the human villin gene was re-sequenced for confirmation. The 12.4-kb regulatory fragment from the mouse villin gene was sequenced in its entirety. An alignment was performed between human (−8.6 kb through intron 1) and mouse (−6.7 kb through intron 1) sequence using the Pustell DNA matrix algorithm on MacVector software. To facilitate the alignment, Alu sequences were identified and removed from the human sequence. Fig. 7A illustrates the considerable sequence conservation in multiple short (20–400 bp) elements of the promoter and first intron. The straight diagonal line indicates that the spacing of conserved elements has also been maintained. Conserved elements in intron 1 lie largely within the region deleted in ΔNheVilLacZ transgenic mice.

Use of the Villin 12.4-kb Regulatory Fragment to Drive Cre Expression in the Intestine of Transgenic Mice—The reproduc-
VilLacZ mice (line 12.4-1) are positive for β-galactosidase and non-expressing cell populations. In adults, crypts of villin are positive for β-galactosidase-negative in ΔNhe line N2 (C) and positive in line 12.4-1 (D). Arrows in C point to the boundary between β-galactosidase expression and non-expressing cell populations. In adults, crypts of ΔNhe line N5 show severely reduced β-galactosidase expression (E), whereas those of line 12.4-1 exhibit robust staining in crypts (F). A similar pattern is noted in the colon (G–J). Pre-crypt cells of the colon of newborn (G) and adult (H) 12.4KbVilLacZ mice are also positive for β-galactosidase. Pre-crypt cells of the colon of newborn ΔNheVilLacZ mice are also positive for β-galactosidase (I), but adults show reduced staining in crypts (J) with a distinct boundary between expressing and non-expressing cells (arrow).

In this study, we describe a 12.4-kb fragment from the mouse villin gene that reliably and reproducibly drives reporter gene expression in the small and large intestines in a pattern that closely resembles that of the endogenous mouse gene. Further, continuous and high level expression of β-galactosidase in 12.4KbVilLacZ mice indicated that this regulatory fragment could be a valuable tool for the generation of conditional genetic mutations. Therefore, we linked the 12.4-kb region to Cre recombinase, generated transgenic founders, and mated these mice (12.4KbVilCre) to the Rosa26 conditional reporter strain, R26R (21). Cre-mediated recombination at loxP sites surrounding a neomycin gene at the Rosa26 locus results in excision of this gene (and its linked polyadenylation sites) and subsequent expression of β-galactosidase.

Intestines from F1 transgenic pups obtained from matings to R26R mice were stained with β-galactosidase to assess the efficacy and specificity of Cre expression. Continuous and high level β-galactosidase expression was seen in the small and large intestines of these pups in five of the seven lines already examined (Fig. 8, A and B). Two other lines show patchy or mottled expression indicative of Cre expression in a subset of crypts (Fig. 8C). Although the continuous expression of Cre in all crypts is required to generate a complete conditional gene modification in the intestine, patchy expression of Cre recombinase will also be of value. Such a line will allow examination of modified and unmodified crypts and villi in the same intestine.

Interestingly, one of the seven Cre lines also exhibits Cre expression in the kidney (data not shown). In an earlier study, one of four founders carrying a 9-kb villin transgene exhibited reporter gene activity in the kidney (8). It is likely, therefore, that these villin regulatory sequences are able to drive expression in the kidney in the context of certain chromosomal environments.

**DISCUSSION**

In this study, we describe a 12.4-kb fragment from the mouse villin gene that reliably and reproducibly drives reporter gene expression in the small and large intestines in a pattern that closely resembles that of the endogenous mouse gene. Further, continuous and high level expression of β-galactosidase in 12.4KbVilLacZ mice indicated that this regulatory fragment could be a valuable tool for the generation of conditional genetic mutations. Therefore, we linked the 12.4-kb region to Cre recombinase, generated transgenic founders, and mated these mice (12.4KbVilCre) to the Rosa26 conditional reporter strain, R26R (21). Cre-mediated recombination at loxP sites surrounding a neomycin gene at the Rosa26 locus results in excision of this gene (and its linked polyadenylation sites) and subsequent expression of β-galactosidase.

Intestines from F1 transgenic pups obtained from matings to R26R mice were stained with β-galactosidase to assess the efficacy and specificity of Cre expression. Continuous and high level β-galactosidase expression was seen in the small and large intestines of these pups in five of the seven lines already examined (Fig. 8, A and B). Two other lines show patchy or mottled expression indicative of Cre expression in a subset of crypts (Fig. 8C). Although the continuous expression of Cre in all crypts is required to generate a complete conditional gene modification in the intestine, patchy expression of Cre recombinase will also be of value. Such a line will allow examination of modified and unmodified crypts and villi in the same intestine.

Interestingly, one of the seven Cre lines also exhibits Cre expression in the kidney (data not shown). In an earlier study, one of four founders carrying a 9-kb villin transgene exhibited reporter gene activity in the kidney (8). It is likely, therefore, that these villin regulatory sequences are able to drive expression in the kidney in the context of certain chromosomal environments.

**FIG. 6.** β-Galactosidase activity is reduced in emerging crypt cells during the formation of intestinal crypts 1 week after birth. A, C, E, I, and J: ΔNheVilLacZ lines. B, D, F, G, and H: 12.4KbVilLacZ lines. A and B, at birth (p0), the intervillus region is continuously stained in small intestine of both ΔNhe line N2 (A) and in line 12.4-1 (B). Seven days later (p7), as initial crypt cells emerge, the nascent cells are β-galactosidase-negative in ΔNhe line N2 (C) and positive in line 12.4-1 (D). Arrows in C point to the boundary between β-galactosidase expression and non-expressing cell populations. In adults, crypts of ΔNhe line N5 show severely reduced β-galactosidase expression (E), whereas those of line 12.4-1 exhibit robust staining in crypts (F). A similar pattern is noted in the colon (G–J). Pre-crypt cells of the colon of newborn (G) and adult (H) 12.4KbVilLacZ mice (line 12.4-1) are positive for β-galactosidase. Pre-crypt cells of the colon of newborn ΔNheVilLacZ mice are also positive for β-galactosidase (I), but adults show reduced staining in crypts (J) with a distinct boundary between expressing and non-expressing cells (arrow).

**FIG. 7.** Identification of phylogenetically conserved elements within intron 1 and the proximal promoter region of the human and mouse villin genes. The entire 12.4-kb murine villin fragment was sequenced, and intron and upstream sequences were aligned with human intron sequences (after removal of Alu repeats). A, the diagonal line generated by this comparison indicates the presence of scattered short homologous elements located in a similar spatial arrangement from −200 bp through the first 3.3 kb of intron 1. B, a schematic diagram of the mouse and human villin genomic loci is shown beneath the homology matrix and aligned vertically with the matrix. The triangles on the human locus demarcate the positions of removed Alu sequences. The gray-shaded box indicates the region containing the majority of homology in the matrix. The 3′ regions of both genomic loci indicated here, including the 3′-most two kilobases of mouse intron 1, and the 3′-most 500 bp of human intron 1, show little or no homology. Likewise, homology is reduced upstream of position −206 (the upstream non-translated exon, UTR, begins at position +1 for both sequences), although several short homologous elements have been found between −206 and −800 (data not shown). It is clear that the NheI deletion (shown as a bold line labeled ΔNhe) removes several homologous elements.
sequences responsible for expression of villin in the stomach are different from those needed for intestinal expression and are not contained within the 12.4-kb regulatory fragment tested.

The data presented here indicate that a multifunctional enhancer (or multiple enhancers) within intron 1 of villin directs expression in discrete areas of the horizontal (duodenum and cecum) and vertical (crypt) axis of the intestine. Thus, different regulatory circuits control villin expression in the duodenum versus the rest of the small intestine, in the cecum versus the rest of the large intestine, and in crypts versus villus tips. This type of region-specific control is reminiscent of the regulation of the Drosophila even-skipped gene, in which several small discrete enhancers control expression of eve in each of the seven different parasegment stripes that comprise its expression domain (22). Region-specific control directed by discrete regulatory elements is also seen in the adenosine deaminase gene (23–24). This ubiquitously expressed gene is controlled in specific tissues by distinct enhancers; T-cell-, placenta-, and duodenum-specific enhancers have been described.

The results presented here differ in two aspects from the results of a previous transgenic study in which villin regulatory sequences were used to drive β-galactosidase (8). First, Pinto et al. reported that deletion of the entire first intron from the 9-kb genomic fragment (their pA3 transgene) resulted in loss of colon expression in three expressing founders (one founder showed no expression at all). In our study, deletion of the 4.2-kb fragment from intron 1, which removes most, but not all of this intron, did not compromise colon expression, although it did increase position effects in accord with the Pinto et al. results (8). Together, these data suggest that elements important for colon-specific expression of villin may lie within intron 1 but outside of the NheI deletion region. Second, the pA3 transgenic founders, which lack intron 1, were reported to express β-galactosidase in the entire intestinal epithelium, including the crypt compartment, although these data were not shown. This result is hard to reconcile with our findings that deletion of a portion of this intron leads to reduced expression in the crypts. This result is not secondary to position effects in our mice, because we observed the same phenotype in multiple offspring of two independently bred lines. Although the different genetic backgrounds of the transgenic lines produced in the two studies could play some modulatory role, the reason for this disparity in results is presently not clear.

Several attributes of the 12.4-kb regulatory fragment suggest that it contains LCR-like activity. First, expression of β-galactosidase was observed in all founders isolated, including two single-copy number lines. Second, expression was continuous in all cells of six out of eight founders, with no evidence of mottled or patchy expression. In the adult intestine, position effect variegation normally presents itself in the form of a mottled expression pattern. This is a direct reflection of the fact that crypts are monoclonal and that stochastic silencing of transgene expression occurs in some proportion of crypts (18). Third, expression was copy number-dependent for all six low copy number lines carrying the 12.4KbVilLacZ construct. Finally, specific regions of the villin 12.4-kb genomic fragment can be implicated in LCR-like function. The region between −6.7 and −3.5 kb appears to play a role, because a 9-kb region of the villin gene (from −3.5 kb through the first intron) showed patchy or heterogeneous expression in 80% of founders studied (8), whereas addition of the three extra kilobases on our construct reduced the proportion of heterogeneously expressing lines to 25%. In addition, deletion of the 4.2-kb region of intron 1 resulted in the apparent loss of copy number dependence and greater variability in transgene expression, indicating that this region also harbors sequences with LCR-like activity. It is possible that the upstream and intronic regions act together to
ensure position independence, similar to what has been observed in the β globin locus (26).

There is one other description of an intestinal LCR. A duodenum-specific enhancer was recently described in the second intron of the human adenosine deaminase gene, but a promoter fragment, including this intron, was unable to drive copy number-dependent expression of a chloramphenicol acetyltransferase reporter in transgenic mice (24). Interestingly, a T-cell-specific enhancer of the human adenosine deaminase gene, located in the first intron, was able to direct copy number-dependent expression of chloramphenicol acetyltransferase in the thymus and in the duodenum when in the presence of a genomic fragment containing the duodenum-specific enhancer (24).

The villin crypt element identified here appears to be necessary for efficient activation of reporter gene expression in the crypt compartment of the entire intestine. Indeed, as this compartment emerges from the intervillus epithelium in the week after birth, an immediate and discrete compartment boundary, which, in differentiated villus tip cells, the 4.2-kb fragment is required for region-specific expression (duodenum and cecum), is visible. It is also of interest that, in differentiated villus tip cells, the 4.2-kb fragment is required for region-specific expression (duodenum and cecum), whereas its activity in crypt cells is not region-specific but similar throughout the large and small intestines. Thus, a regulatory circuitry, to which the villin gene responds, may be held in common by crypts along the entire intestinal tract. Alternatively, there may be more than one genetic program for establishment of villin expression in the crypt compartment across the length of the intestine, and multiple cis elements within the 4.2-kb NheI fragment may exist that are responsive to each of these programs. A more precise identification of one or more specific cis response elements for such regulatory networks will facilitate the elucidation of the trans factors involved in this pathway. For future studies, it will be important to determine whether this 4.2-kb fragment (or subfragments thereof) is both necessary and sufficient to confer crypt-specific expression on a heterologous transgene.

In summary, we have defined a 12.4-kb genomic region of the villin gene that is capable of driving transgene expression in all cells of the crypt and tip and along the entire large and small intestines. This genomic region also contains LCR-like sequences that suppress position effect variegation. We have already used this regulatory fragment to derive mice that express Cre recombinase efficiently in the intestinal compartment, a tool that will be extremely valuable for the creation of conditional genetic mutations. Our identification of multiple regulatory activities within the first intron of the villin gene, and the recognition of significant evolutionary conservation within this region, will aid in the future investigation of the cis and trans factors involved in a number of spatial patterning pathways important in intestinal organogenesis.

Acknowledgments—B. B. M. is grateful for support from the Training Program in Organogenesis. D. L. G. acknowledges support from the Roger McDermitt Research Fund, and from a Cancer Innovation Grant, both from the University of Michigan Comprehensive Cancer Center. We are grateful to Christine Babcock for help with animal breeding and genotyping. We also thank the University of Michigan Transgenic Animal Core for transgenic mouse production and Dr. Stuart Orkin and members of his laboratory for assistance with 12.4KbVilCre construct assembly.

REFERENCES
1. Bieker, J. J. (1998) Curr. Opin. Hematol. 5, 145–150
2. Orkin, S. H. (1996) Eur. J. Biochem. 231, 271–281
3. Zaret, K. S. (2000) Mech. Dev. 92, 83–88
4. Braunstein, E. M., Qiao, X. T., Madison, B., Pinson, K., Dunbar, L., and Gumucio, D. L. (2002) Dev. Dyn. 224, 90–102
5. Ezzell, R. M., Chafel, M. M., and Matsudaira, P. T. (1989) Development 106, 407–419
6. Maunoury, R., Rohine, S., Pringault, E., Leonard, N., Gaillard, J. A., and Louvard, D. (1992) Development 115, 717–728
7. Maunoury, R., Rohine, S., Pringault, E., Huet, C., Guenet, J. L., Gaillard, J. A., and Louvard, D. (1998) EMBO J. 7, 3321–3329
8. Pinto, D., Rohine, S., Saisser, F., El Mariou, F. E., and Louvard, D. (1999) J. Biol. Chem. 274, 6476–6482
9. Rohine, S., Sahuquillo-Merino, C., Louvard, D., and Pringault, E. (1993) J. Biol. Chem. 268, 11426–11434
10. Pinson, K. I., Dunbar, L., Samuelson, L., and Gumucio, D. L. (1998) Dev. Dyn. 211, 109–121
11. Dillon, N., and Grosveld, P. (1993) Trends Genet. 9, 134–137
12. Felsenfeld, G. (1992) Nature 355, 219–224
13. Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Milioni, A., Jones, M., and Kioussis, D. (1996) Science 271, 1123–1125
14. Reitman, M., Lee, E., Westphal, H., and Felsenfeld, G. (1990) Nature 348, 749–752
15. Su, Y., Liebhaber, S. A., and Cooke, N. E. (2000) J. Biol. Chem. 275, 7902–7909
16. Downs, K. M., and Davies, T. (1993) Development 118, 1255–1266
17. Kaufman, M. H. (1994) The Atlas of Mouse Development, pp. 95–248, Academic Press, London
18. Roth, K. A., Hermiston, M. L., and Gordon, J. I. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 9407–9411
19. Rubin, D. C., Ong, D. E., and Gordon, J. I. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 1278–1282
20. Crossman, M. W., Hauft, S. M., and Gordon, J. I. (1994) J. Cell Biol. 126, 1547–1564
21. Soriano, P. (1999) Nat. Genet. 21, 70–71
22. Sackerson, C., Pujikasa, M., and Goto, T. (1999) Dev. Biol. 211, 39–52
23. Dusing, M. R., Brickner, A. G., Lowe, S. Y., Cohen, M. B., and Wigington, D. A. (2000) Am. J. Physiol. 279, G186–G199
24. Dusing, M. R., Brickner, A. G., Thomas, M. B., and Wigington, D. A. (1997) J. Biol. Chem. 272, 26634–26642
25. Dusing, M. R., Florence, E. A., and Wigington, D. A. (2001) J. Biol. Chem. 276, 14434–14442
26. Stamatoyannopoulos, J. A., Clegg, C. H., and Li, Q. (1997) Mol. Cell. Biol. 17, 240–247