MOLECULAR AND CELLULAR BIOLOGY, Jan. 1995, p. 407–414 Vol. 15, No. 1
0270-7306/95/$04.00+0

Expression of the Chicken β-Globin Gene Cluster in Mice: Correct Developmental Expression and Distributed Control

MARK M. MASON,[1] ERIC LEE,[2] HEINER WESTPHAL,[2] AND MARC REITMAN[1]*

Diabetes Branch, National Institute of Diabetes and Digestive and Kidney Diseases,[1] and Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, Bethesda, Maryland 20892

Received 14 July 1994/Returned for modification 15 September 1994/Accepted 21 September 1994

To investigate the regulation of gene clusters, we introduced the entire chicken β-globin cluster into mice. This 35-kb region includes the four globin genes (ρ-β^H-β^A-ε), the four upstream hypersensitive sites, and the intergenic β^A/ε enhancer. The chicken globins are not arranged in order of developmental expression, which is unlike the case for the human β-globin cluster, in which gene order plays a role in the regulation of globin expression. Mice carrying the chicken cluster expressed the transgenes with the same developmental patterns as seen in the chicken. Therefore, stage-specific erythroid transcriptional milieus existed before the divergence of birds and mammals and have been conserved since then. Mice bearing the complete cluster except for a deletion removing the β^A/ε enhancer displayed markedly reduced expression of the β^H, β^A, and ε genes with efficient (but variable) ρ expression. Mice carrying the four genes and β^A/ε enhancer but without the upstream hypersensitive sites showed reduced expression of ρ, β^H, and β^A, with variable expression of ε. We conclude that (i) all of the genes (except possibly ρ) are under the control of both the upstream hypersensitive sites and the enhancer, (ii) the influence of the control elements can extend beyond the nearest active gene, (iii) a single element (the enhancer) can influence more than one gene in a single developmental stage, (iv) the enhancer can work bidirectionally, and (v) neither the upstream sites (as a group) nor the enhancer showed developmental stage specificity. Thus, the regulation of this cluster is achieved by interaction of two distinct control regions with each of the globin genes.

Genes are not distributed at random throughout the genome. Specialized domains of chromosomes, such as telomeres and centromeres, exclude genes. Even within gene-replete areas of the genome, genes are nonrandomly distributed, frequently being clustered in small regions of DNA. These clusters often consist of related genes, having arisen by gene duplication. Although some aspects of gene cluster regulation are unique to clusters (such as the effect of gene order on gene expression), it is likely that regulation of genes in clusters occurs by the same mechanisms that regulate individual genes. However, by studying gene cluster regulation, one can separate early events influencing the whole cluster (e.g., chromatin opening) from later ones affecting individual genes (e.g., transcription). By studying the expression of genes in intact and mutated clusters, one can identify the control regions that mediate regulatory events.

The human β-globin locus is a well-studied vertebrate gene cluster. It consists of five genes arranged in order of developmental expression (5′-ε-^{G}γ-^{A}γ-δ-β-3′) and five upstream hypersensitive sites in ~65 kb of DNA (40). The upstream hypersensitive sites (21, 45) are essential for globin expression at a high level and in a copy number-dependent manner in transgenic mice (23). These sites, referred to as a locus control region (LCR), contribute to early replication in erythroid cells and to chromatin opening (20). Exactly which of these properties define the term LCR is not established. The ability to confer copy number dependence is redundantly distributed among the upstream hypersensitive sites (reference 22 and references therein) and can be dissociated from high-level expression (42).

* Corresponding author. Mailing address: National Institutes of Health, Diabetes Branch, Bldg. 10, Room 8S-239, Bethesda, MD 20892-1770. Phone: (301) 496-6090. Fax: (301) 402-0573. Electronic mail address: mlr@helix.nih.gov.

Developmental stage-specific regulation of the human globin cluster is postulated to involve successive physical interactions between the upstream sites and the local regulatory regions of the individual genes (reviewed in references 9, 11, 14, 15, and 43). Formation of a particular LCR-gene interaction depends on the proteins associated with the upstream sites and the genes, the distance between the LCR and gene, and the lack of competing interactions (24, 28, 32).

The mechanisms for choosing individual genes for expression in the chicken β-globin cluster are interesting for several reasons. First, unlike the human cluster, the chicken genes (5′-ρ-β^H-β^A-ε-3′) are not arranged in order of developmental expression—the embryonic genes (ρ and ε) flank the fetal (β^H) and fetal/adult (β^A) genes. Second, the avian and mammalian clusters evolved independently from a single β-globin gene in the common ancestor (10), and therefore different gene choice regulatory mechanisms may have evolved. Third, the chicken cluster contains a strong enhancer between β^A and ε (6, 26), while in the mammalian clusters, the strong enhancer is found upstream of the genes. The chicken β^A/ε enhancer mediates copy number-dependent expression of β^A in mice (36). This enhancer interacts with the β^A promoter through GATA-1 and a stage-specific factor (5). The enhancer is not sufficient to open chromatin but requires the cooperation of a promoter (37). The ρ and β^A promoters are intrinsically stage specific (26, 30). In contrast, developmental regulation of ε is proposed to occur via competition between the ε and β^A promoters for the enhancer (7, 18). Upstream of the cluster are three erythroid cell-specific hypersensitive sites, two of which (5′HS2 and 5′HS3) have moderate enhancer activity (1). Farthest upstream is a constitutive hypersensitive site (5′HS4) that can act as an insulator or boundary element (8).

In this study, we have produced transgenic mice carrying either the complete chicken β-globin cluster or the cluster with specific deletions, either of the upstream hypersensitive sites or

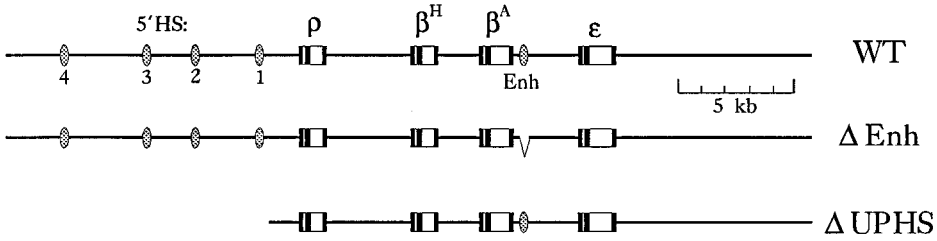

FIG. 1. Constructs used to produce transgenic mice. The wild-type chicken locus (WT) is shown, with the four β-like globin genes (ρ, β^H, β^A, and ε), the β^A/ε intergenic enhancer (Enh), and the upstream hypersensitive sites (5′HS1 to -4) indicated. The ρ and ε genes are expressed from embryonic day 2 to 5, β^H is expressed from embryonic day 6 through hatching, and β^A is transcribed from embryonic day 6 through adulthood. Also depicted are the ΔEnh and ΔUPHS constructs.

of the β^A/ε enhancer. The expression and developmental regulation of all four of the chicken globin genes were examined in these lines.

MATERIALS AND METHODS

Chicken β-globin cluster constructions. The cosmid pCos8β containing ~35 kb of the chicken β-globin cluster in pWE15 was provided by M. Barton (it is the same as pCos9β in reference 4; Fig. 1). The ΔEnh construct was made by deleting the β^A/ε enhancer (478 bases between *Hha*I sites) from pCos8β, using RecA-assisted restriction endonuclease cleavage (RARE [17]). Briefly, cosmid DNA (5 μg) was incubated with RecA (5 μg; U.S. Biochemical) and 300 ng of oligonucleotides (coding strand bases 17608 to 17667 and 18086 to 18145 in GenBank locus CHKHBBRE (accession number L17432 [35]) in the appropriate buffer at 37°C for 10 min. The reaction mixtures were maintained at 37°C, while *S*-adenosylmethionine (to 250 μM) and *Hha*I methylase (40 U) were added. After methylation (37°C, 30 min) and inactivation of the RecA and methylase (65°C, 20 min), the DNA was digested with *Hha*I (80 U, 37°C, 60 min), religated, packaged (Gigapack II packaging extract; Stratagene), and transformed into SURE cells (Stratagene), and the desired clone (pMM2) was isolated.

The ΔUPHS construct (~24 kb) was made by using the RARE technique to create a unique *Pvu*I site at the *Aci*I site 1,304 bp upstream of the ρ cap site. An oligonucleotide (6814 to 6873 in CHKHBBRE) was used to protect the *Aci*I site in pCos8β from *Sss*I methylase. After digestion with *Aci*I, a *Pvu*I linker (CGGGCGATCGCC) was inserted, the DNA was packaged and used to transform cells, and the desired clone (pMM6) was isolated.

Production of transgenic mice. DNA for injection was obtained free of vector by *Not*I (wild-type construct [WT] and ΔEnh) or *Not*I and *Pvu*I (ΔUPHS) digestion followed by separation using sedimentation in an NaCl gradient (2). Transgenic FVB/N mice (41) were produced by microinjecting the DNA at a low concentration (1 ng/μl) into the male pronuclei (27) and screened by PCR on tail DNA (37). To eliminate the confounding effects of mosaicism and multiple integration sites, only progeny mice were used in our experiments. Male transgenic progeny were bred with superovulated wild-type FVB/N females to produce transgenic embryos of known gestational age. Day 0 is the morning that the vaginal plug was observed, or about 0.5 day postcoitus.

DNA hybridization. Southern blot analysis was performed as previously described (34). Six different unique-sequence probes were used: I100 (a 555-bp *Bgl*II fragment found upstream of ρ, 5848 to 6403 in CHKHBBRE), I95 (a ~500-bp *Xba*I-*Hin*dIII fragment from pCBG21.6 [46] found downstream of ε), and I92, I99, I142, and I144 (34). Copy number was estimated from Southern and slot blots by comparison with spiked mouse DNA standards on the same blot (36). The intensity was measured with a PhosphorImager and is estimated to be within ±25% of the actual value.

RNA isolation and RNase protection assay. RNA was isolated from yolk sac, liver, and blood samples, using RNA-stat 60 (Tel-Test). An RNA protection assay was developed because of limited sensitivity and high backgrounds with use of primer extension. The β^H/β^A probe detects the first exons of β^H and β^A. Plasmid p1109 is the 275-bp genomic *Bam*HI-*Ava*II fragment including β^H exon 1 (12861 to 13136 in CHKHBBRE) inserted between the *Bam*HI and *Eco*RV sites of pBSIISK− (Stratagene). The β^H/β^A antisense probe (368 nucleotides [nt]) was made by using T7 polymerase on *Bam*HI-cut p1109. The major products from this probe are the expected 142-nt protected fragment from β^H mRNA and the 94-nt fragment from β^A mRNA.

Since the first exons of ρ and ε are identical, the second exon was used to distinguish these RNAs. Plasmid p1098 is the 375-bp genomic *Pvu*II-*Eag*I fragment including ε exon 2 (20465 to 20840 in CHKHBBRE) inserted between the *Eag*I and *Sma*I sites of pBSIISK−. The ρ/ε antisense probe (430 nt) was made by using T3 polymerase on *Hin*dIII-cut p1098. The observed protected products obtained by using this probe are the expected 222-nt fragment from ε mRNA and a ~115-nt fragment from ρ mRNA. The observed ρ product is longer than the expected 86 nt and is explained by the RNases not cleaving at two single-base mismatches and/or sequence polymorphisms at these positions. Extra bands, close in size to the β^H product, were sometimes observed with the ρ/ε probe in samples containing ρ and ε RNAs and with the β^{maj} probe in samples containing β^{maj} RNA. Therefore, quantitation of β^H was done in the absence of these probes.

pSP64Mζ (mouse ζ-globin) and pSP64Mβ134 (mouse β^{maj}-globin) were used as described previously (3) except that pSP64Mβ134 was linearized with *Sfa*NI, generating a probe of 100 nt and a protected product of 76 nt.

RNase protection assays were performed as described elsewhere (2) except that hybridization conditions (70°C in 90% formamide–400 mM NaCl–20 mM Tris HCl [pH 7.5]–1 mM EDTA–0.1% sodium dodecyl sulfate) and RNase digestion (37°C, 60 min) were adjusted to increase the assay's specificity. After denaturing gel electrophoresis, the signals were quantitated with a PhosphorImager. To facilitate comparison between experiments, a standard mixture of RNAs from 5-day (0.5 μg) and 11-day (0.5 μg) embryonic chicken blood and nontransgenic adult mouse blood (0.65 μg) was assayed with each set of samples.

RESULTS

Generation of transgenic mice. To study the developmental regulation of chicken β-globin gene expression, we made transgenic mice carrying (i) the intact chicken cluster (WT), (ii) the cluster minus the enhancer found between the β^A and ε genes (ΔEnh), or (iii) the cluster minus the upstream hypersensitive sites (ΔUPHS). The WT construct is a 35-kb genomic fragment containing the four genes and all of the known control elements with ~3 kb of upstream and ~7 kb of downstream flanking DNA (Fig. 1). RARE (17) was used to create the ΔEnh construct via deletion of a 478-bp fragment containing the β^A/ε enhancer. The ΔUPHS construct is missing ~11 kb of upstream DNA containing the four upstream hypersensitive sites and ending 1304 bp 5′ of the ρ cap site. We generated three lines of mice carrying the WT fragment, four lines containing the ΔEnh construct, and five lines bearing the ΔUPHS construct. The ΔUPHS-4 and ΔUPHS-5 lines are F_1 progeny from a single founder that carried two independently segregating transgene integration sites. The F_1 ΔUPHS-4 and ΔUPHS-5 mice did not produce any transgenic F_2 progeny.

The chicken cluster is intact in the transgenic lines. Southern blot analysis established that the transgenes were not rearranged. Figure 2 shows a blot of tail DNA hybridized to a mixture of six unique probes, which detect fragments that account for 83% of the WT and ΔEnh constructs and 60% of the ΔUPHS DNA. The bands were of the expected size and relative intensity, making it unlikely that rearrangements of the transgenes had occurred. ΔUPHS-5 had a minor extra band, suggesting that most copies were intact but that one copy was rearranged. Other blots probed with the complete WT cosmid and blots of DNA digested with *Bam*HI and *Asp*718 confirmed that the transgenes were not rearranged (data not shown). Transgene copy number was estimated from the Southern blots and slot blots. These results are presented below. It is notable that six of the lines were single-copy integrants.

Expression of globin genes in chicken blood. Before examining chicken β-globin expression in transgenic mice, we quantitated the pattern of RNA levels in the chicken. An RNase

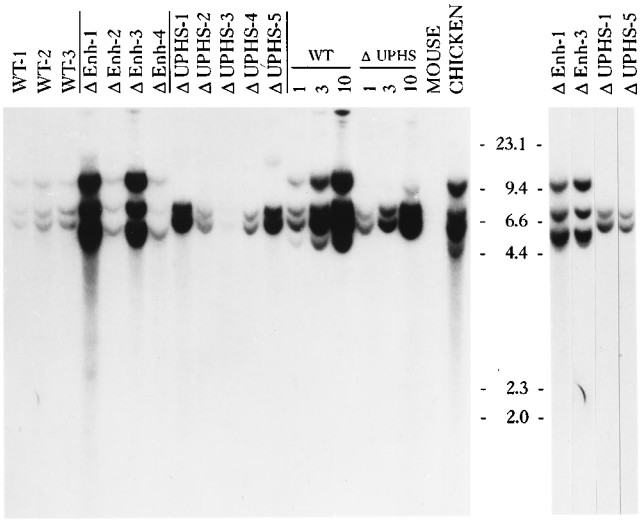

FIG. 2. Southern blot of transgenic mouse DNA. DNA from progeny of each transgenic line was digested with *Eco*RI, electrophoresed, blotted, and hybridized to a mixture of six probes that spanned the WT insert (see Materials and Methods). Controls include chicken DNA (CHICKEN), nontransgenic mouse DNA (MOUSE), and nontransgenic mouse DNA spiked with WT or ΔUPHS insert DNA at 1, 3, and 10 copies per diploid genome. The left panel is a 16-h exposure. A 4-h exposure of the ΔEnh-1, ΔEnh-3, ΔUPHS-1, and ΔUPHS-5 lanes is shown on the right. The sizes (in kilobases) of the standards are indicated. The WT-1 and ΔUPHS-3 lanes were underloaded, as revealed by ethidium bromide staining.

protection assay using one probe to detect ρ and ε and another for β^H and β^A was developed and used to measure these mRNAs in embryonic blood (Fig. 3). The ρ- and ε-globins were the only β-like RNAs in 5-day blood, and they were present at similar levels. The ρ and ε RNAs decreased to undetectable levels by day 15. Expression of β^A was seen by day 7 and continued through day 15. The β^H expression pattern in ovo mimicked that of β^A but at a level ~8-fold lower. Others have shown that after hatching, β^H is turned off, and β^A is the only β-like globin in adult chickens (38). These results validate the RNase protection assay and augment the previous determinations of chicken globin RNA levels (references 29 and 30 and references therein).

Expression of β^A-globin in adult transgenic mice. As our first test for transgene expression, we measured RNA levels in blood from adult mice. All of the WT, ΔEnh, and ΔUPHS lines expressed the β^A-globin transgene (Fig. 4). None of the lines expressed detectable ρ-, β^H-, or ε-globin RNA. Thus, the expression pattern of the four chicken globins in blood from adult mice is the same as that in adult chickens. β^A expression per transgene copy was quantitatively similar in the three WT lines, averaging 16% of the mouse β^{maj} gene level (Fig. 4B). In contrast, in the ΔEnh mice, expression per copy varied over a 24-fold range and averaged only 0.33% of the β^{maj}-globin level. These data confirm that the enhancer is necessary for both copy number dependence and a high level of β^A expression. In the five ΔUPHS lines, β^A expression averaged 2.7% of β^{maj} expression and showed little variation in expression per copy. This is the same expression level as in mice carrying a 4.4-kb fragment including the β^A gene and its 3′ enhancer (36). The lack of effect of the 20 kb of additional DNA contained in ΔUPHS (including three more genes and numerous regulatory elements) suggests that this DNA does not contain regulatory elements that act on β^A. Comparing β^A expression in the WT and ΔUPHS lines demonstrates that removal of the upstream hypersensitive sites did not affect copy number dependence but lowered β^A expression by ~6-fold.

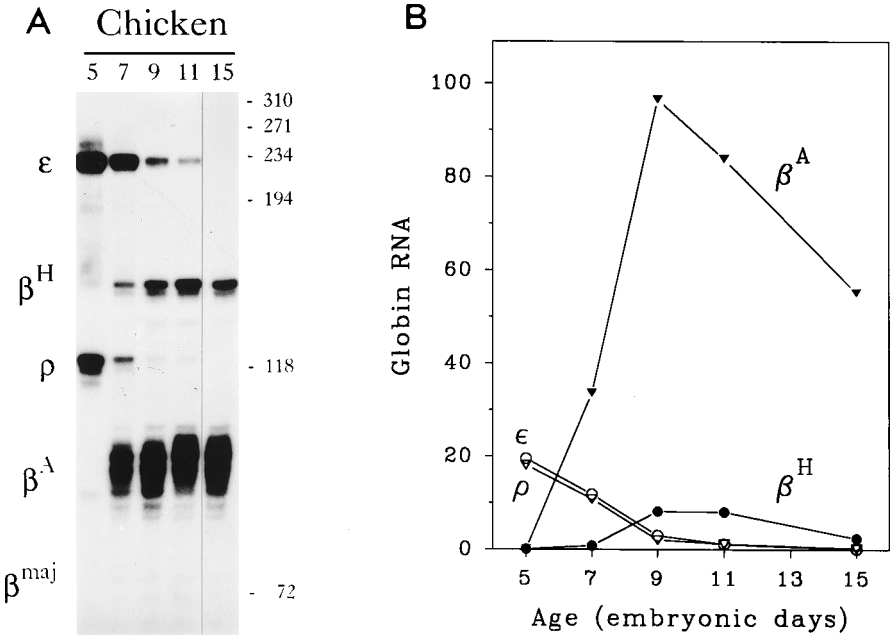

FIG. 3. Globin RNA in the chicken embryo. (A) Chicken blood RNA (3 μg) from 5-, 7-, 9-, 11-, and 15-day embryos was assayed for the four chicken β-like globin genes and for mouse β^{maj}-globin by RNase protection (see Materials and Methods). Shown on the left are the mobilities of the ε, β^H, ρ, β^A, and β^{maj}-globin protected products. At the right are positions of the single-stranded DNA size standards (in nucleotides). (B) Data from two experiments (one shown in panel A) were quantitated, averaged, and corrected for the specific activity of protected products. The resulting RNA levels are in moles, plotted on a single arbitrary scale for all four globins: (○), ρ (▽), β^A (▼), and β^H (●).

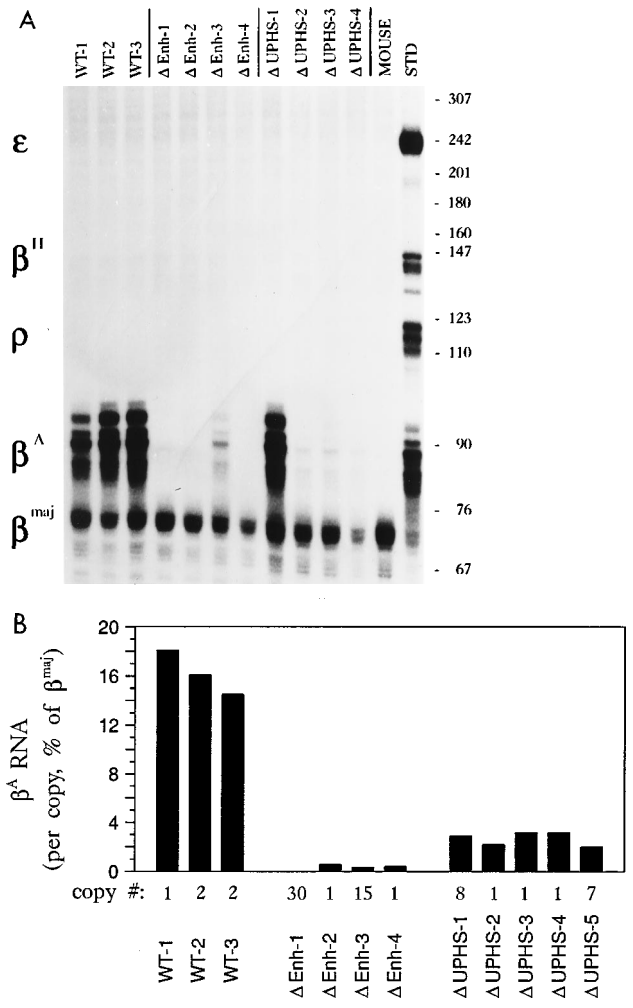

FIG. 4. Transgene expression in blood from adult mice. (A) RNA (3 μg) was isolated from blood of adult mice of the indicated lines, and RNase protection was performed with the ρ/ε, β^H/β^A, and mouse β^{maj} probes. The product mobilities are indicated at the left. A nontransgenic mouse control (MOUSE) and a standard RNA mixture (STD; see Materials and Methods) are also shown. At the right are positions of single-stranded DNA size standards (in nucleotides). (B) β^A-Globin expression in adult blood. The amount of β^A RNA in adult blood was quantitated and normalized to the β^{maj} internal control. The normalized β^A data are expressed per transgene copy, which is shown at the bottom. ΔEnh-1 expression was measured at 0.024% of β^{maj} per copy. Data are means of two to four assays except those for ΔUPHS-4 and ΔUPHS-5, which are from single determinations.

Developmental stage and tissue specificity of transgene expression. The developmental expression pattern was examined in detail in four of the transgenic lines. Murine embryonic globin expression occurs in primitive erythrocytes made in 9- to 11-day yolk sac, fetal globins are transcribed in definitive cells made in 12- to 16-day liver, and adult globin RNA is found in reticulocytes in blood of adult animals (49). The cells in embryonic blood change from predominantly primitive erythrocytes at day 12 to definitive erythrocytes by day 16. All four transgenes in the WT-1 line were expressed in a pattern analogous to that seen in the chicken (Fig. 5). RNAs for ρ and ε, but not β^H or β^A, were detected in 9- to 11-day yolk sac. Twelve- to sixteen-day fetal liver contained β^A transcripts. The small amounts of ρ and ε in 12-day liver are attributed to contamination by circulating blood. In blood, the ρ and ε levels decreased markedly from days 12 to 16 while β^A increased. In more sensitive assays, β^H RNA was detected in 16-day fetal blood but not in blood from adult mice (data not shown and Fig. 6). From these data, we conclude that in the WT-1 line, ρ and ε are embryonic globins, β^H is a fetal globin, and β^A is a fetal/adult globin.

Transgene expression in the WT-2 line was examined in the same manner as the WT-1 line. The expression patterns of ρ, β^H, β^A, and ε were qualitatively and quantitatively similar in the two lines except that in the WT-2 line, a β^H signal was seen in the 14- and 16-day liver and in 14- and 16-day blood (data not shown). The β^H signal was strongest in the 16-day blood.

The transcription pattern for the ΔEnh-2 line is shown in Fig. 5. The ρ transgene had an embryonic profile and was expressed as efficiently as in the WT-1 line. In contrast, ε RNA was undetectable in ΔEnh-2 embryos. The β^A RNA levels were considerably reduced but continued to show a fetal/adult expression pattern. Although not visible in Fig. 5, a more sensitive assay detected β^H RNA in 16-day blood (Fig. 6).

In the ΔUPHS-3 line, β^A was expressed at a low level in 16-day and adult blood but was not detected in other erythroid tissues (9-, 10-, and 11-day yolk sac, 12-, 14-, and 16-day liver, and 12- and 14-day blood; data not shown). No ρ, β^H, or ε RNA was detected in these tissues in this line (data not shown).

Tissue specificity of transgene expression was examined in three lines (WT-1, ΔEnh-4, and ΔUPHS-2). Multiple tissues (uterus, brain, muscle, kidney, liver, and heart) from adult mice were assayed for transgene RNA. In no case was ρ, β^H, or ε RNA detected. The small amount of β^A detected was attributed to blood contamination since the β^A/β^{maj} ratio was the same as in blood and spleen (data not shown).

Expression of ε and ρ in 12-day transgenic mice. We used 12-day blood to measure embryonic globin expression since the developmental series data displayed the largest signal in this sample. Both ε and ρ were expressed efficiently in the WT lines (Fig. 7), with ε expression slightly greater than that of ρ (59 versus 19% of each gene's levels in 5-day chicken blood RNA). In the ΔEnh lines, expression of ε was severely decreased (~30-fold) while expression of ρ averaged 122% of the level in WT lines. ρ-Globin expression in the ΔEnh lines varied in correlation with copy number. The lines expressing ρ efficiently (ΔEnh-2 and ΔEnh-4) carried one copy of the transgene, while ΔEnh-1 and ΔEnh-3, with 30 and 15 copies, showed reduced expression per copy (but similar total expression). Possible explanations for the ρ data are (i) ρ expression is saturated, resulting in less efficient per-copy expression in the higher-copy-number lines, as seen in mice carrying human ζ-globin genes (39); (ii) in the multicopy lines, the control elements interact in an improper configuration, resulting in decreased ρ expression or expression of only one of the ρ copies (such as an end copy); and (iii) copy number dependence is not expected, since the enhancer is required for copy number-dependent expression. We cannot distinguish among these possibilities.

In the ΔUPHS lines, ρ expression was at least eightfold less than in the WT lines. ε expression in the single-copy lines (ΔUPHS-2 and ΔUPHS-3) was also markedly reduced. In contrast, ΔUPHS-1, with eight copies, showed a level of ε expression per copy comparable to levels in the WT lines.

Thus, the expression of ρ is much more (or completely) dependent on the upstream hypersensitive sites than on the β^A/ε enhancer. Optimal ε expression requires both the upstream sites and the enhancer, perhaps with a greater contribution from the enhancer.

Expression of β^H in 16-day transgenic mice. Fetal blood from 16-day embryos was assayed for β^H RNA (Fig. 6). In the three WT lines, β^H expression averaged 5.4% of its level in

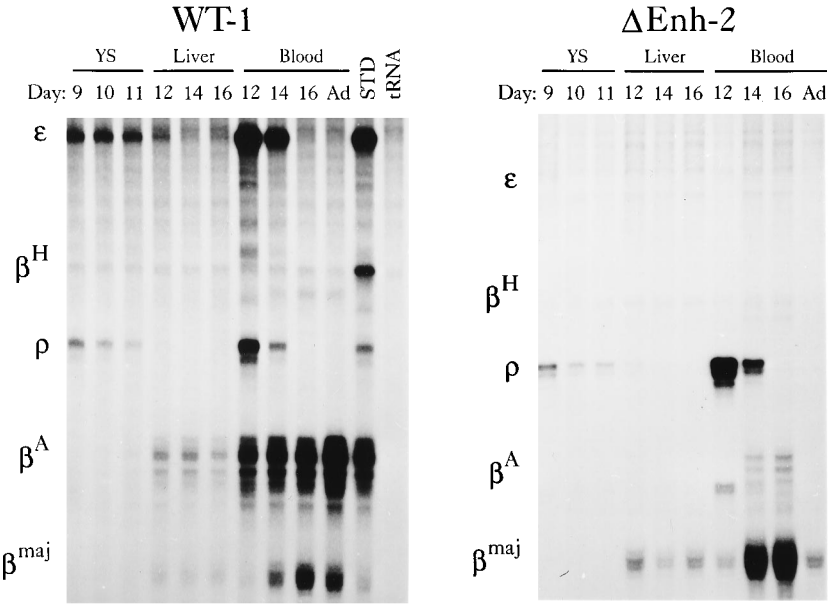

FIG. 5. Developmental patterns of expression in the WT-1 and ΔEnh-2 lines. An RNase protection assay using WT-1 RNA from 9-, 10-, and 11-day embryonic yolk sac (YS; 5 μg of RNA), 12-, 14-, and 16-day fetal liver (3 μg), 12-, 14-, and 16-day fetal blood (1 μg), and adult blood (1 μg) is on the left. tRNA and standard RNA mixture (STD) controls are so indicated. The mobilities of the ε, β^H, ρ, β^A, and β^{maj}-globin protected products and single-stranded DNA size standards (in nucleotides) are also shown. On the right is an RNase protection experiment using RNA from the ΔEnh-2 line. Details are the same as for WT-1 except that 3 μg of yolk sac RNA was tested. Both assays used the ρ/ε, β^H/β^A, and mouse β^{maj} probes.

11-day chicken blood and 2.3% of the level of the β^A gene in the same sample. β^H expression was markedly reduced by deletion of the β^A/ε enhancer, with an average decrease of 35-fold in a comparison of the ΔEnh and WT lines. However, in the two single-copy ΔEnh lines, β^H expression was still ~2% of β^A expression, while in the 30- and 15-copy lines, it was 12 and 36% of β^A expression. Thus, β^H expression, like that of ρ, was different in the multicopy ΔEnh lines than in the single-copy lines. The direction of the β^H effect (more expression in multicopy) was opposite that observed with ρ (less expression in multicopy). Deletion of the upstream sites reduced β^H expression at least 50-fold (to <1% of β^A expression) in all three ΔUPHS lines. We conclude that β^H expression in mice is even less efficient than in chickens, but both the upstream sites and the enhancer contribute to the expression seen.

DISCUSSION

Correct developmental stage-specific expression of the four chicken β-like globins in mice. Our results demonstrate that the chicken β-like globins are developmentally regulated in transgenic mice and that the pattern of expression is strikingly similar to that seen in the chicken. In the terminology of mammalian globin expression, ρ and ε are embryonic globins, β^H is fetal and β^A is fetal/adult. These results demonstrate that embryo-, fetus-, and adult-specific erythroid transcriptional milieus were established before, and have been conserved since, the divergence of birds and mammals. Consistent with our results, a *Xenopus* tadpole globin functioned as an embryonic globin in a transgenic mouse line, suggesting that at least two erythroid transcriptional environments existed at the amphibian-mammalian divergence, which predates the avian-mammalian divergence (12). In our experiments, stage specificity was preserved in the lines missing the enhancer or the upstream hypersensitive sites. These data suggest that stage specificity is determined by the gene promoters.

The last common ancestor of birds and mammals had only one β-globin gene but had two α-like genes, the embryonic ζ and the fetal/adult α (10, 33). Thus, the α-like globins are candidates for the genes that provided the selective advantage for the establishment of stage-specific erythroid transcription. Since the β-like globins duplicated later, they evolved to use the preexisting stage-specific transcription environments and did not play a role in the establishment of these milieus.

In the WT lines, transgene expression was reasonably efficient but not quite up to the levels seen in the chicken, with RNA levels averaging 19% for ρ, 13% for β^A, 5% for β^H, and 59% for ε of the level in chicken blood. The lower levels suggest that the heterologous system only partially mimics transcription in the chicken. Consistent with this explanation, the amino acid sequence of chicken GATA-1 is very different from that of mammalian GATA-1 outside of the DNA-binding region (see reference 44). Thus, mouse GATA-1 bound to the chicken regulatory regions might make less efficient contacts with other proteins. Divergence of other genes could also contribute to the lower levels of expression. However, in spite of these quantitative differences, the information gained from studying chicken globins should be applicable to understanding mammalian stage-specific erythroid transcription.

LCR activity in the chicken β-globin cluster. We have shown previously that a 4.4-kb region including the β^A gene and its 3′ enhancer has LCR activity (using copy number-dependent expression in transgenic mice as the assay for LCR activity [36]). We also showed that the enhancer by itself, even in multicopy lines, did not open chromatin (another frequently used definition of LCR activity). Our interpretation is that the enhancer must interact with other regulatory elements (probably the promoter) to open chromatin (37). A complementary result is that more than two copies of human 5′HS2 were required for detectable β-globin transgene expression (13). Combining these observations leads to the speculation that LCR activity is

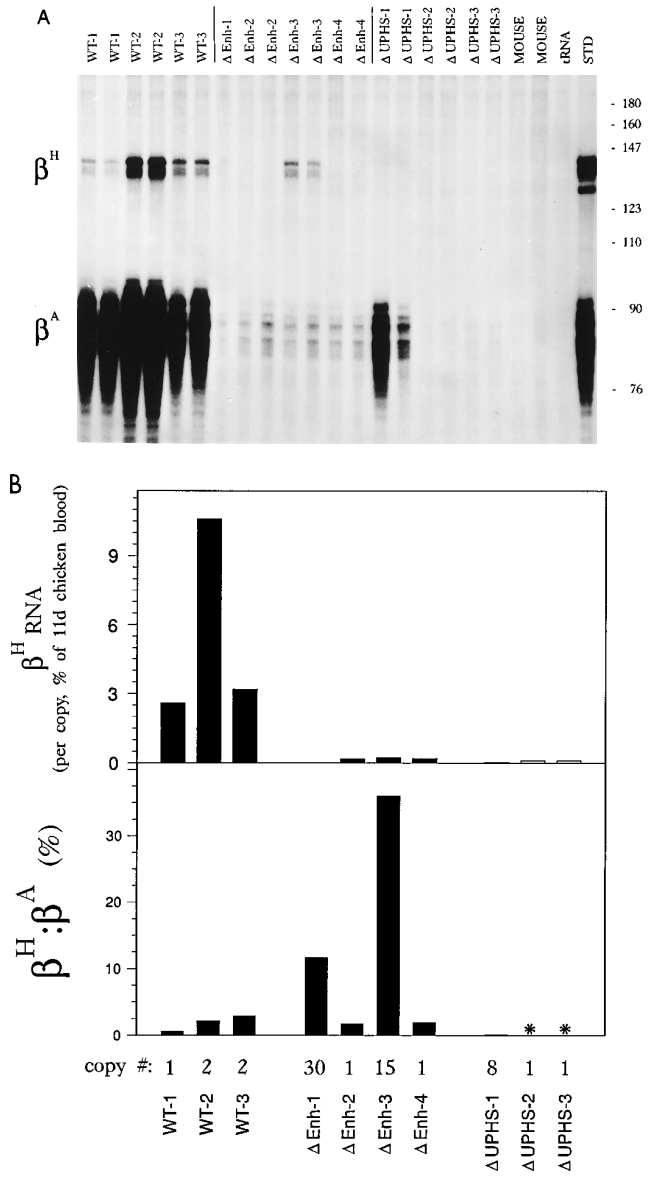

FIG. 6. Transgene expression in blood from 16-day embryos. (A) RNase protection was performed on RNA (3 μg) from two embryos of each line (except ΔEnh-1), using the $\beta^H\beta^A$ probe. Controls and labeling are the same as in Fig. 5. (B) Quantitation of β^H-globin expression. In the top panel, the β^H RNA level is presented, per copy number, as a percentage of the β^H RNA level in 11-day (11d) chicken blood. In the ΔUPHS-2 and ΔUPHS-3 samples, no signal was detected, and so the limit of detection is shown as an open bar. In the lower panel, the β^H level is presented as a percentage of the β^A signal in the same sample.

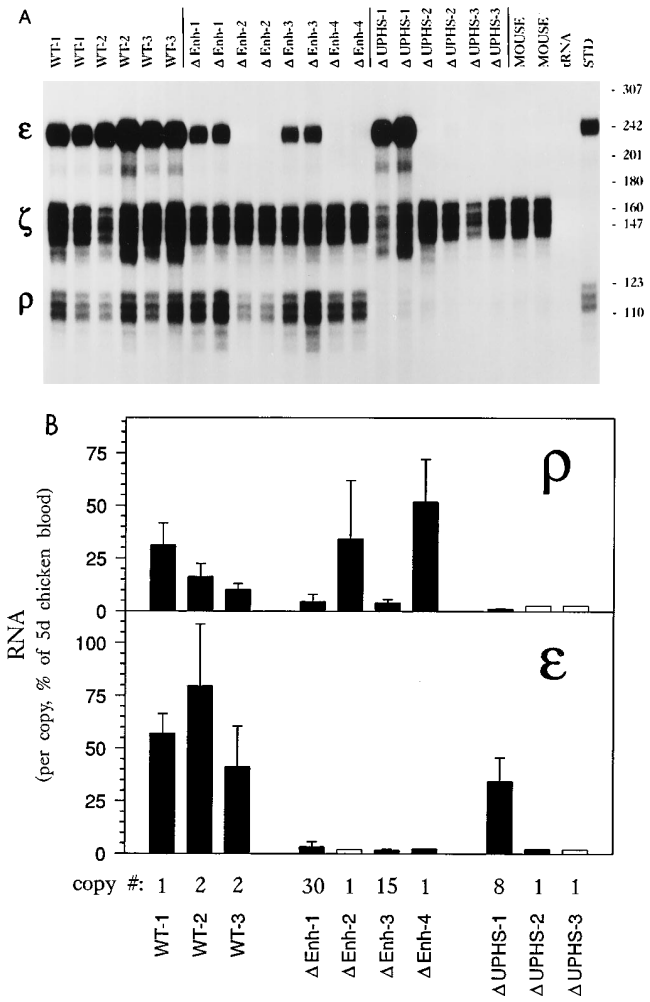

FIG. 7. Transgene expression in blood from 12-day embryos. (A) RNase protection was performed on RNA (2 μg) from two embryos of each indicated line, using the ρ/ε and mouse ζ probes. Controls and labeling are the same as in Fig. 5. (B) ρ- and ε-globin expression in 12-day blood. The amounts of ε and ρ RNAs were quantitated and normalized to the ζ internal control. Data are expressed per transgene copy as a percentage of the amount of ρ or ε found in 5-day (5d) chicken blood. Data are means of two to six determinations, and error bars represent 1 standard deviation. When an RNA was not detected, the limit of detection (twice the background) is shown as an open bar. Although shown as undetectable, ε expression in ΔEnh-2 and ΔUPHS-3 was probably present at about the limit of detection.

a property resulting from the interaction of a promoter with multiple other elements.

As already noted, the location of LCR activity inside the chicken cluster is different from that of the mammalian globin clusters. A region essential for the chicken LCR activity, the β^A/ε enhancer, is postulated to be derived from upstream sites (35). Do the chicken upstream sites, themselves, contain LCR activity? If chromatin opening is used to define LCR activity, they do. As a rule, transcribed genes have open, hypersensitive promoters, without canonical nucleosomes. The four ΔEnh lines all transcribe their transgenes, so chromatin opening has occurred. Since a construct missing the enhancer and the upstream sites did not transcribe (36), it is likely that the upstream sites contributed to the transcription. The most stringent definition of LCR activity is copy number-dependent expression of a linked gene in transgenic mice. By this definition, neither the upstream sites nor the enhancer, by themselves, have complete LCR activity since the expression per copy varies for some of the genes (e.g., ρ for the ΔEnh construct and ε for the ΔUPHS construct).

Distributed control: optimal transgene expression requires both the β^A/ε enhancer and the upstream hypersensitive sites. Figure 8 summarizes the contribution of the upstream sites and enhancer to the expression of the four chicken globins. We conclude that (i) all of the genes, with the possible exception of ρ, are under the control of both the upstream hypersensitive sites and the enhancer; (ii) each control element's influence can extend past the nearest active gene to a more distant one (discussed in detail below); (iii) a single element can influence

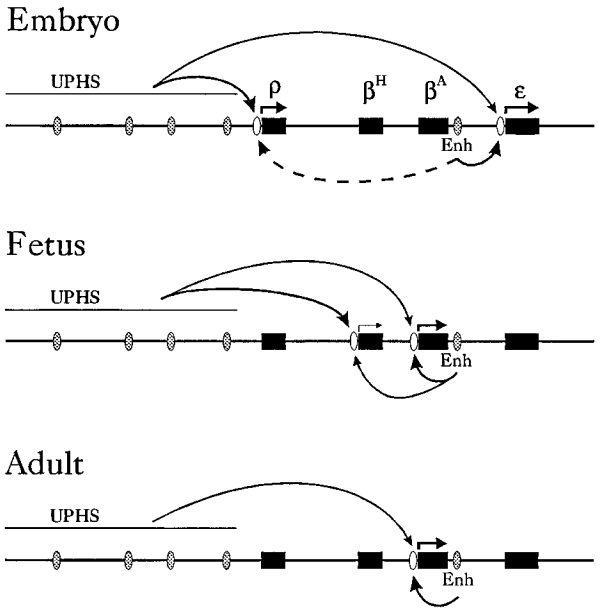

FIG. 8. Regulation of the chicken β-globin gene cluster. The pattern of influence of the upstream hypersensitive sites (UPHS) and the β^A/ε enhancer (Enh) on each globin gene's expression is indicated for each developmental stage. Stage-specific hypersensitive sites at the promoters are shown as open ovals; panerythroid and constitutive sites are shaded. Transcription is indicated above the gene. Thicker arrows indicate stronger effects. The dashed arrow indicates uncertainty in interpretation of the data. See text for details.

more than one gene in a single developmental stage (the enhancer affects both β^H and β^A in fetal cells); (iv) the control elements can work bidirectionally (the enhancer works on both β^A and ε in the single copy lines, extending earlier conclusions [7, 18, 31]); and (v) neither the upstream sites, as a group, nor the enhancer showed developmental stage specificity (note, however, that demonstration of developmental specificity in the human cluster required examination of individual sites [22]). From the directional competition model of the human β-globin regulation, one might predict that the upstream sites control the ρ, β^H, and β^A genes while the enhancer controls only ε. This is clearly not the case.

Our working model for the regulation of the chicken β-globin gene cluster is that the whole cluster (33 kb) acquires an open chromatin configuration before, or coincidently with, the start of transcription (25, 47). This opening requires interaction between the regulatory regions (enhancer and upstream sites) and the promoters that are active in the specific lineage. Previous experiments using transient expression suggested that the ε promoter was not stage specific and that stage-specific ε transcription resulted from competition between the β^A and ε promoters for the β^A/ε enhancer (7, 18). However, recent studies in transgenic mice did not confirm this hypothesis (19). The expression pattern of the β^H and β^A genes in our transgenic lines also argues against promoter competition for the enhancer. Deletion of the upstream sites or enhancer reduced expression of both β^H and β^A, whereas if directional competition were occurring, the expression of the distal gene should be reduced much more severely. The slightly greater influence of each region on the nearer gene may be a proximity effect (32). The expression pattern of ρ and ε in the transgenic lines also argues against promoter competition (the upstream sites did influence ε expression). The human β-globin gene, for which competition is invoked as a regulatory mechanism, may be a special case. In the human cluster, active genes are found between the upstream hypersensitive sites and the β gene in the stages that do not express β. Thus, the β gene is exposed to the upstream regions only in those cells in which it is expressed, and there is no selective pressure for stage specificity of the human β promoter. In contrast, the chicken cluster, with two sets of regulatory elements, has ample selective pressure for maintaining strict developmental stage specificity in its promoters.

How do the distant sites regulate the genes? Conventional wisdom postulates that a physical interaction occurs between the distant control regions and the promoters (a looping mechanism [16]). Evidence in favor of this hypothesis has been very difficult to obtain. One model proposes binary interactions between specific upstream hypersensitive sites and individual promoters (14). Our result that both the upstream region and the enhancer affect the expression of individual genes is evidence against a strict binary interaction model. Generalizing to the human cluster, we predict that multiple elements will be important for the expression of the $^G\gamma$ gene, and the same elements will also control the $^A\gamma$ genes. Similarly, both the δ and β genes are expected to be controlled by overlapping sets of elements.

We have demonstrated that the transcription factor environments controlling the developmental stage-specific expression of the β-globin genes are highly conserved between birds and mammals. We also showed that, unlike the case for the human β-globin cluster, the regulation of the chicken β-globin cluster genes is mediated by two distinct control regions in addition to the globin genes' promoters. It is intriguing that the chicken control regions are separated by three genes and many kilobases of DNA. Elucidating the details of the multiple interactions between the genes and the distant control regions will further our understanding of gene cluster regulation.

ACKNOWLEDGMENTS

We thank M. Barton, B. Emerson, and M. Baron for clones, L. Ferrin for advice on RARE, and J. Chung, A. Dean, G. Felsenfeld, A. Ginsberg, J. Grasso, and S. Taylor for comments on the manuscript.

M.R. is a Lucille P. Markey Scholar, and this work was supported in part by a grant from the Lucille P. Markey Charitable Trust.

REFERENCES

1. **Abruzzo, L. V., and M. Reitman.** Enhancer activity of upstream hypersensitive site 2 of the chicken β-globin cluster is mediated by GATA sites. J. Biol. Chem., in press.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1993. Current protocols in molecular biology. Wiley, New York.
3. **Baron, M. H., and T. Maniatis.** 1986. Rapid reprogramming of globin gene expression in transient heterokaryons. Cell **46:**591–602.
4. **Barton, M. C., M. F. Hoekstra, and B. M. Emerson.** 1990. Site-directed, recombination-mediated mutagenesis of a complex gene locus. Nucleic Acids Res. **18:**7349–7355.
5. **Barton, M. C., N. Madani, and B. M. Emerson.** 1993. The erythroid protein of cGATA-1 functions with a stage-specific factor to activate transcription of chromatin-assembled β-globin genes. Genes Dev. **7:**1796–1809.
6. **Choi, O.-R., and J. D. Engel.** 1986. A 3′ enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. Nature (London) **323:**731–734.
7. **Choi, O.-R., and J. D. Engel.** 1988. Developmental regulation of β-globin gene switching. Cell **55:**17–26.
8. **Chung, J. H., M. Whiteley, and G. Felsenfeld.** 1993. A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell **74:**505–514.
9. **Crossley, M., and S. H. Orkin.** 1993. Regulation of the β-globin locus. Curr. Biol. **3:**232–237.
10. **Czelusinak, J., M. Goodman, D. Hewett-Emmett, M. L. Weiss, P. J. Venta, and R. E. Tashian.** 1982. Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature (London) **298:**297–300.
11. **Dillon, N., and F. Grosveld.** 1993. Transcriptional regulation of multigene

loci: multilevel control. Trends Genet. **9:**134–137.
12. **Dillon, N., G. Kollias, F. Grosveld, and J. G. Williams.** 1991. Expression of adult and tadpole specific globin genes from *Xenopus laevis* in transgenic mice. Nucleic Acids Res. **19:**6227–6230.
13. **Ellis, J., D. Talbot, N. Dillon, and F. Grosveld.** 1993. Synthetic human β-globin 5′HS2 constructs function as locus control regions only in multicopy transgene concatemers. EMBO J. **12:**127–134.
14. **Engel, J. D.** 1993. Developmental regulation of human β-globin gene transcription: a switch of loyalties? Trends Genet. **9:**304–309.
15. **Evans, T., G. Felsenfeld, and M. Reitman.** 1990. Control of globin gene transcription. Annu. Rev. Cell Biol. **6:**95–124.
16. **Felsenfeld, G.** 1992. Chromatin as an essential part of the transcriptional mechanism. Nature (London) **355:**219–224.
17. **Ferrin, L. J., and R. D. Camerini-Otero.** 1991. Selective cleavage of human DNA: RecA-assisted restriction endonuclease (RARE) cleavage. Science **254:**1494–1497.
18. **Foley, K. P., and J. D. Engel.** 1992. Individual stage selector element mutations lead to reciprocal changes in β- vs. ε-globin gene transcription: genetic confirmation of promoter competition during globin gene switching. Genes Dev. **6:**730–744.
19. **Foley, K. P., S. Pruzina, J. D. Winick, J. D. Engel, F. Grosveld, and P. Fraser.** 1994. The chicken β/ε-globin gene enhancer directs autonomously regulated, high level expression of the chicken ε-globin gene in transgenic mice. Proc. Natl. Acad. Sci. USA **91:**7252–7256.
20. **Forrester, W. C., E. Epner, M. C. Driscoll, T. Enver, M. Brice, T. Papayannopoulou, and M. Groudine.** 1990. A deletion of the human β-globin locus activation region causes a major alteration in chromatin structure and replication across the entire β-globin locus. Genes Dev. **4:**1637–1649.
21. **Forrester, W. C., S. Takegawa, T. Papayannopoulou, G. Stamatoyannopoulos, and M. Groudine.** 1987. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. Nucleic Acids Res. **15:**10159–10177.
22. **Fraser, P., S. Pruzina, M. Antoniou, and F. Grosveld.** 1993. Each hypersensitive site of the human β-globin control region confers a different developmental pattern of expression on the globin genes. Genes Dev. **7:**106–113.
23. **Grosveld, F., G. Blom van Assendelft, D. Greaves, and G. Kollias.** 1987. Position-independent, high level expression of the human β-globin gene in transgenic mice. Cell **51:**975–985.
24. **Hanscombe, O., D. Whyatt, P. Fraser, N. Yannoutsos, D. Greaves, N. Dillon, and F. Grosveld.** 1991. Importance of globin gene order for correct developmental expression. Genes Dev. **5:**1387–1394.
25. **Hebbes, T. R., A. L. Clayton, A. W. Thorne, and C. Crane-Robinson.** 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken β-globin chromosomal domain. EMBO J. **13:**1823–1830.
26. **Hesse, J. E., J. M. Nickol, M. R. Lieber, and G. Felsenfeld.** 1986. Regulated gene expression in transfected primary chicken erythrocytes. Proc. Natl. Acad. Sci. USA **83:**4312–4316.
27. **Hogan, B., F. Costantini, and E. Lacy.** 1986. Manipulating the mouse embryo. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Kim, C. G., E. M. Epner, W. C. Forrester, and M. Groudine.** 1992. Inactivation of the human β-globin gene by targeted insertion into the β-globin locus control region. Genes Dev. **6:**928–938.
29. **Lois, R., and H. G. Martinson.** 1989. Chicken globin gene transcription is cell lineage specific during the time of the switch. Biochemistry **28:**2281–2287.
30. **Minie, M., T. Kimura, and G. Felsenfeld.** 1992. The developmental switch in embryonic ρ-globin expression is correlated with erythroid lineage-specific differences in transcription factor levels. Development **115:**1149–1164.
31. **Nickol, J. M., and G. Felsenfeld.** 1988. Bidirectional control of the chicken β- and ε-globin genes by a shared enhancer. Proc. Natl. Acad. Sci. USA **85:**2548–2552.
32. **Peterson, K. R., and G. Stamatoyannopoulos.** 1993. Role of gene order in developmental control of human γ- and β-globin gene expression. Mol. Cell. Biol. **13:**4836–4843.
33. **Proudfoot, N. J., A. Gil, and T. Maniatis.** 1982. The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell **31:**553–563.
34. **Reitman, M., and G. Felsenfeld.** 1990. Developmental regulation of topoisomerase II sites and DNase I-hypersensitive sites in the chicken β-globin locus. Mol. Cell. Biol. **10:**2774–2786.
35. **Reitman, M., J. A. Grasso, R. Blumenthal, and P. Lewit.** 1993. Primary sequence, evolution and repetitive elements of the *G. gallus* (chicken) β-globin cluster. Genomics **18:**616–626.
36. **Reitman, M., E. Lee, H. Westphal, and G. Felsenfeld.** 1990. Site-independent expression of the chicken β^A-globin gene in transgenic mice. Nature (London) **348:**749–752.
37. **Reitman, M., E. Lee, H. Westphal, and G. Felsenfeld.** 1993. An enhancer/locus control region is not sufficient to open chromatin. Mol. Cell. Biol. **13:**3990–3998.
38. **Roninson, I. B., and V. M. Ingram.** 1983. Expression and partial DNA sequence of the chicken β^H-globin gene. J. Biol. Chem. **258:**802–809.
39. **Sharpe, J. A., D. J. Wells, E. Whitelaw, P. Vyas, D. R. Higgs, and W. G. Wood.** 1993. Analysis of the human α-globin gene cluster in transgenic mice. Proc. Natl. Acad. Sci. USA **90:**11262–11266.
40. **Stamatoyannopoulos, G., and A. W. Nienhuis.** 1994. Hemoglobin switching, p. 107–155. *In* G. Stamatoyannopoulos, A. W. Nienhuis, P. W. Majerus, and H. Varmus (ed.), The molecular basis of blood diseases. W. B. Saunders, Philadelphia.
41. **Taketo, M., A. C. Schroeder, L. E. Mobraaten, K. B. Gunning, G. Hanten, R. R. Fox, T. H. Roderick, C. L. Stewart, F. Lilly, C. T. Hansen, and P. A. Overbeek.** 1991. FVB/N: an inbred mouse strain preferable for transgenic analyses. Proc. Natl. Acad. Sci. USA **88:**2065–2069.
42. **Talbot, D., and F. Grosveld.** 1991. The 5′HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J. **10:**1391–1398.
43. **Townes, T. M., and R. R. Behringer.** 1990. Human globin locus activation region (LAR): role in temporal regulation. Trends Genet. **6:**219–223.
44. **Trainor, C. D., T. Evans, G. Felsenfeld, and M. Boguski.** 1990. Structure and evolution of a human erythroid transcription factor. Nature (London) **343:**92–96.
45. **Tuan, D., W. Solomon, Q. Li, and I. M. London.** 1985. The "beta-like-globin" gene domain in human erythroid cells. Proc. Natl. Acad. Sci. USA **82:**6384–6388.
46. **Villeponteau, B., G. M. Landes, M. J. Prankratz, and H. G. Martinson.** 1982. The chicken β-globin gene region. J. Biol. Chem. **257:**11015–11023.
47. **Weintraub, H., and M. Groudine.** 1976. Chromosomal subunits in active genes have an altered conformation. Science **193:**848–856.
48. **Whitelaw, E., S.-F. Tsai, P. Hogben, and S. H. Orkin.** 1990. Regulated expression of globin chains and the erythroid transcription factor GATA-1 during erythropoiesis in the developing mouse. Mol. Cell. Biol. **10:**6596–6606.