

Developmental regulation of a complete 70-kb human β -globin locus in transgenic mice

John Strouboulis, Niall Dillon, and Frank Grosveld

Laboratory of Gene Structure and Expression, National Institute for Medical Research, London, NW7 1AA, UK

We have used a linker-based ligation strategy to combine two 35-kb cosmid inserts from the human β -globin locus into one linear fragment containing the entire locus. This 70-kb fragment was introduced into transgenic mice by microinjection of fertilized eggs. Southern blot analysis showed that a single complete transgene locus can be introduced into the germ line with high efficiency. Analysis of the expression patterns of the locus during development shows that the ϵ -globin gene behaves as a purely embryonic gene, the γ -globin gene as an embryonic and early fetal gene, and the β -globin gene as a fetal adult gene. Quantitation of expression showed that the levels of transcription of the ϵ - and γ -globin genes are reversed relative to their mouse homologs but that the total output of the human and mouse loci is constant during development. These results suggest that multiple changes in DNA sequences and transcription factor balance must have occurred for the human γ -globin gene to have evolved into a fetal gene.

[Key Words: β -Globin locus; transgenic mice; developmentally regulated gene expression]

Received May 11, 1992; revised version accepted July 20, 1992.

The human β -globin multigene locus contains five functional genes that are expressed in a tissue- and developmental-specific manner. The genes are arranged in the order in which they are expressed during development; ϵ is the most 5' gene and is expressed in the yolk sac, followed by the γ^G and γ^A genes, which are expressed in the fetal liver. The δ and β genes are at the 3' end and are expressed in adult bone marrow, with >98% of the total resulting from the β gene (see Fig. 1, below; for review, see Collins and Weissman 1984). Humans therefore contain a separate set of genes (the γ genes) that are expressed at the fetal stage in contrast to most species, including the mouse, which do not have a separate fetal gene. The entire cluster of genes is regulated by the locus control region (LCR; Grosveld et al. 1987), which is characterized by a set of developmentally stable DNase hypersensitive sites, distributed over a region of 20 kb, 5' to the ϵ -globin gene (Tuan et al. 1985; Forrester et al. 1987; Grosveld et al. 1987). Deletions of these sites observed in a subset of β -thalassemias result in an inactive globin locus (Kioussis et al. 1983; Driscoll et al. 1989; Forrester et al. 1990). Addition of the LCR results in high-level, copy number-dependent expression of linked globin or heterologous genes in erythroid cells in transgenic mice. This expression is independent of the site of integration of the transgene construct in the host genome (Grosveld et al. 1987). A number of deletion/mapping experiments have shown that this property is specified by 200- to 300-bp core sequences at each of the hypersensitive sites (HS1-HS4, Fig. 1; for review, see Dillon et al. 1991).

The complete locus, including the LCR, spans a distance of 70 kb. The location of the LCR at the 5' end of the locus means that there are very great differences in the distances between the individual genes and the LCR. The ϵ gene, for example, is located only 11 kb from HS2 compared with >50 kb for the β gene. This ability to activate genes at very long range is one of the most interesting features of the LCR and has not been studied directly owing to technical limitations of existing cloning methods.

The human β -globin locus is also a prototype system for studying developmental regulation in multigene loci. A large body of experimental evidence has been accumulated in which a variety of single-gene and multigene constructs have been analyzed in transgenic mice. Both the ϵ and γ genes have been found to be expressed embryonically when linked to the LCR in small constructs, and expression was silenced in adult stages (Raich et al. 1990; Shih et al. 1990; Dillon and Grosveld 1991). Interestingly, expression of the γ genes was found to extend into the early fetal stage, a surprising observation as mice do not have separate fetal globin genes. Analysis of β expression using constructs containing both γ and β genes showed that the order of the genes relative to the LCR is capable of strongly affecting the developmental regulation of β with embryonic silencing observed only when β is in the distal position (Hanscombe et al. 1991).

The results summarized above have allowed the formulation of several models for globin gene switching (Townes and Behringer 1990; Dillon and Grosveld 1991;

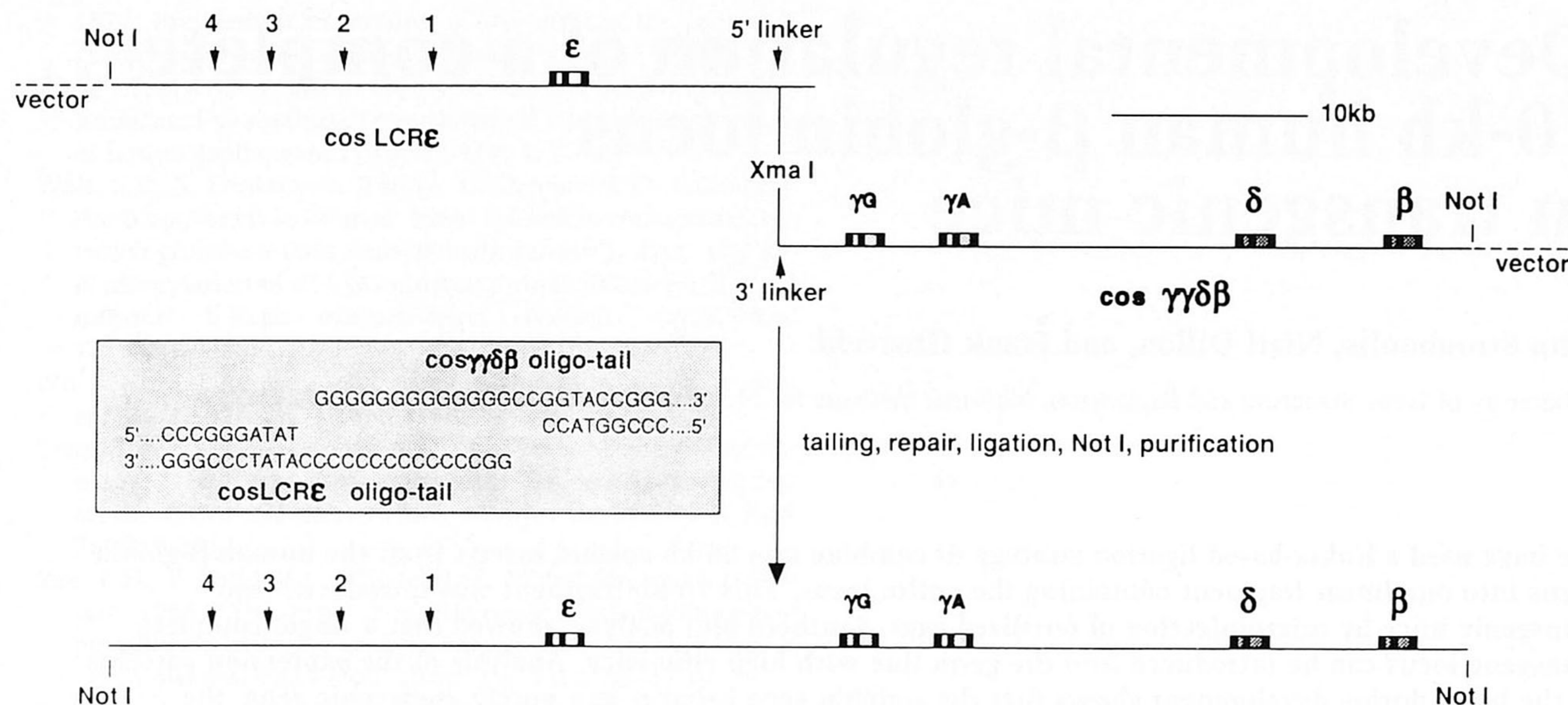


Figure 1. Construction of the human β -globin locus. CosLCRE and cos $\gamma\delta\beta$ were combined by providing a linker (*inset*) at a common *Xma*I restriction site. Arrows 1–4 indicate the positions of the DNase I hypersensitive sites in the LCR (Tuan et al. 1985; Forrester et al. 1987; Grosveld et al. 1987).

Hanscombe et al. 1991; Stamatoyannopoulos 1991). All of these, however, were formulated in the absence of a "normal" baseline, against which the expression of mutated constructs could be compared. Because mice do not have fetal globin genes, there is no way of knowing whether a particular pattern of expression is a feature of the construct being tested, the interaction of the genes with the nonhomologous mouse system, or a combination of both of these parameters. The only way to resolve this problem is to introduce a complete single-copy locus into transgenic mice and to determine the pattern of expression of the genes during murine development. We therefore decided to introduce the entire β -globin locus into transgenic mice by linking cosmids *in vitro* before microinjection into fertilized eggs. In this paper we show that a single copy of the entire human β -globin gene locus can be introduced into the mouse genome efficiently with the surprising result that human γ -globin gene expression is not only observed in the early fetal mouse liver but that it is also the earliest and most abundantly expressed transgene during the embryonic period.

Results

The strategy for the construction of the complete globin locus is shown in Figure 1. Two cosmid clones were modified to allow linkage at a restriction site present in a region of overlap (for details, see J. Strouboulis, N. Dillon, and F. Grosveld, *in prep.*). Linkers containing a 15-bp poly(dG-dC) stretch, either preceded or followed by a unique restriction site, were inserted into the *Xma*I site of each cosmid. In addition, a unique *Not*I site was added to the 5' or 3' border to allow removal of the vector sequences. The two resulting cosmids, cosLCRE and cos $\gamma\delta\beta$, were linearized with *Nae*I and *Sma*I, respec-

tively, and treated with T4 DNA polymerase in the presence of dATP and TTP. Under these conditions, the exonuclease activity of T4 polymerase generates overlapping single stranded ends at high efficiency. The tail of cos $\gamma\delta\beta$ was partially repaired with dCTP to create a perfect fit with the complementary cosLCRE tail (Fig. 1). Vector sequences were removed by cleavage with *Not*I followed by salt gradient sedimentation (see Materials and methods). The cosmid inserts were annealed and ligated, and the mixture of fragments was fractionated by preparative pulse-field gel electrophoresis. The ligated 70-kb fragment was found to be the major species (>70%) and was electroeluted from the gel, exchanged into injection buffer by passage through a Sephadex G50 column, and microinjected into fertilized mouse eggs. Forty-eight mice were born and screened for the presence of human sequences by Southern blot analysis of tail DNA, using probes specific for the 5' and 3' regions of the β -globin locus. A total of nine mice were transgenic for human globin sequences, four of which were positive for both ends of the locus (nos. 2, 13, 15, and 72). The cosmid joining region was also intact in these four mice, as demonstrated by polymerase chain reaction (PCR) amplification across the junction at which the cosmids had been annealed (not shown). The other five mice appeared to have deletions and/or rearrangements and were not analyzed further.

To confirm that the locus had not been rearranged or deleted in other regions, Southern blots were hybridized to the complete starting cosmids, LCRE and $\gamma\delta\beta$, as well as to probes specific for the 5' and 3' ends of the locus. Whole cosmid patterns of hybridization showed each *Eco*RI (Fig. 2) and *Hind*III (not shown) fragment of the locus to be present in all four mice, including the 7-kb *Eco*RI joining fragment (solid triangle, Fig. 2). End-frag-

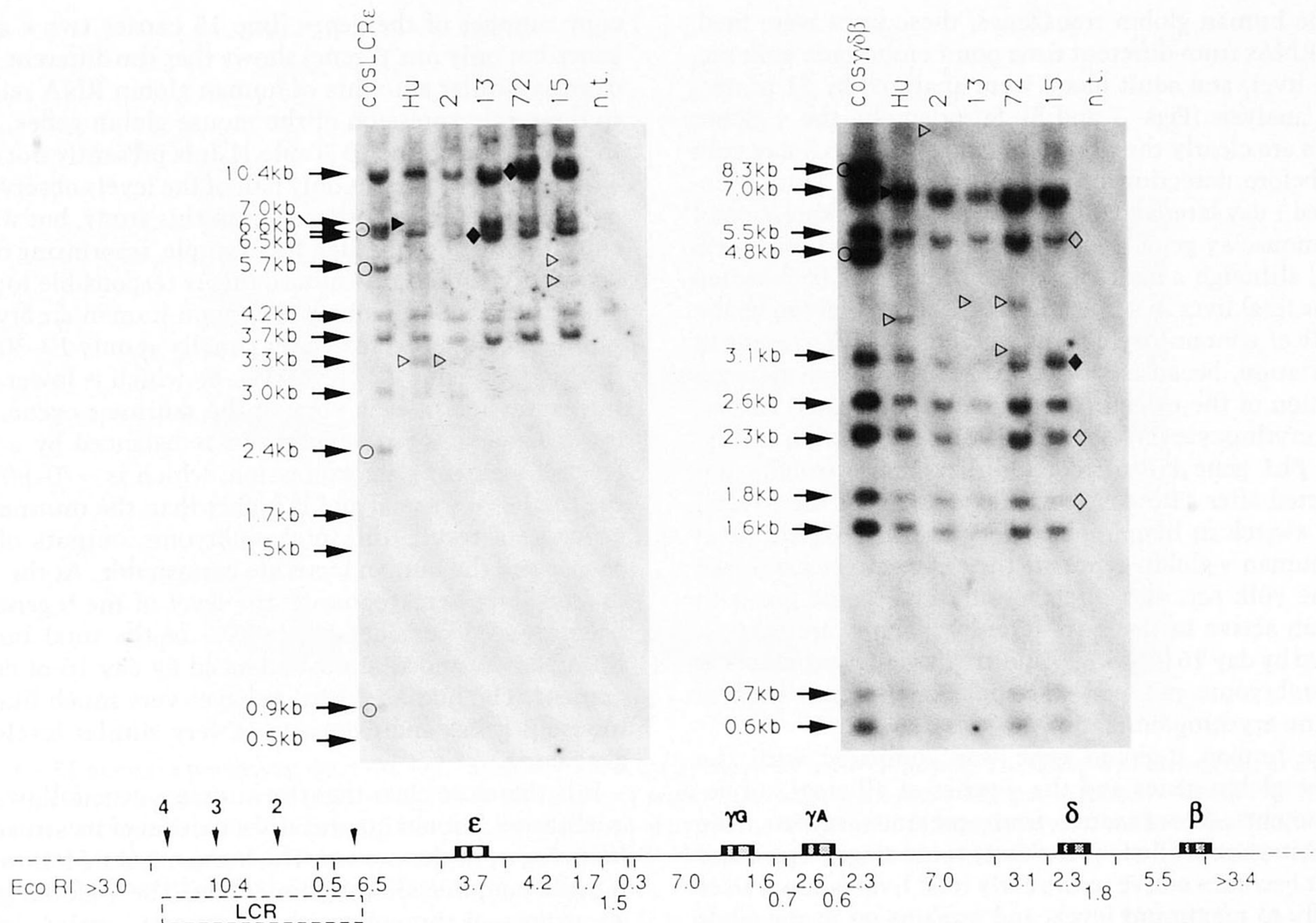


Figure 2. Southern blots of the transgenic human loci in mouse lines 2, 13, 71, and 15 digested with *Eco*RI. Control lanes contain the starting cosmids cosLCRe, cos $\gamma\gamma\delta\beta$ (diluted in mouse genomic DNA), and human (Hu) DNA. (Left) Blot hybridized with the complete cosLCRe; (right) blot hybridized with cos $\gamma\gamma\delta\beta$. Both probes were preannealed with total human and mouse DNA to compete out repetitive sequences. Arrows show the sizes of the various bands. The *Eco*RI map of the human locus is shown at bottom. (Δ) End fragments; (\blacklozenge) end fragments that fall in a doublet. (\circ) Vector DNA fragments not present in the microinjected 70-kb fragment. (\blacktriangle ; lane labeled Hu) The joining fragment that has been reconstituted in the transgenic DNA lanes by the joining of the cosmids (see map). In the cos $\gamma\gamma\delta\beta$ blot, this fragment falls within a doublet also indicated (\blacktriangle). (\diamond ; right panel) Fragments with a lower intensity relative to the other bands. The 3.0-kb band in all lanes of the left panel is occasionally observed and appears to be a cross-hybridizing mouse fragment. The 1.7-kb and 1.5-kb bands mapping between ϵ and γ (left panel) are always very weak because they consist almost entirely of repetitive DNA, which is competed under these hybridization conditions (Taramelli et al. 1986).

ment hybridization (denoted by open triangles and solid diamonds, Fig. 2) showed that the locus was present in a single copy in three of the mice (nos. 2, 13, and 72). Although line 72 contains two rather than one 3'-end fragment, the absence of a second β -gene fragment suggested that one of these end fragments is all that remains of a second locus and that it is not ligated to the other completely intact locus. This was confirmed by Southern blotting of a *Hind*III digest. For line 15, the 5'-end probe hybridization showed two fragments (open triangles, Fig. 2), consistent with the integration of two copies of the locus, but only one 3'-end fragment could be detected (solid diamond, Fig. 2) and fragments from the 3' part of the locus had a lower signal intensity (open diamonds, Fig. 2). This suggested that one copy carried a deletion, which was subsequently confirmed and mapped to the 2.3-kb *Eco*RI fragment at the 5' end of the δ gene, deleting the complete δ and β genes (not shown). The full locus and the deleted locus are again not ligated

to each other, as demonstrated by the absence of a fragment hybridizing to both the 5'- and 3'-end probes (not shown). However, they appear to be closely linked because both were present in all the offspring. These types of separate but linked integrations have also been reported for small fragments (Singh et al. 1991). Correct linkage of all the internal *Eco*RI fragments for all of the lines was confirmed by detection of the normal overlapping fragments on blots of *Hind*III digests. In the case of line 15, this also showed the presence of two separate end fragments, one for each locus, at the expected 5' position. In addition it showed one 3'-end fragment at the expected position and one 3' fragment that maps to the δ gene position. We therefore conclude that in all four lines one copy of the human locus has integrated unrearranged, whereas line 15 has two-thirds of a second locus. This results in a line that carries two copies of the ϵ and γ genes but only a single copy of the δ and β genes. To establish the developmental pattern of expression

of the human globin transgenes, these lines were bred, and RNAs from different time point embryonic yolk sac, fetal liver, and adult blood were analyzed by S1 protection analysis (Figs. 3 and 5). Surprisingly, the γ -globin genes are clearly the first to be expressed in 8.5-day yolk sac before detection of ϵ gene expression. This is observed 1 day later and follows the pattern of expression of the mouse $\epsilon\gamma$ gene. $\epsilon\gamma$ and ϵ are switched off after 11.5 days, although a faint signal can sometimes be detected in the fetal liver in some animals. This is likely to be the result of contaminating embryonic cells still present in circulation, because previous work has shown that expression of the $\epsilon\gamma$ gene is confined to nucleated embryonic erythrocytes (Wong et al. 1983). The murine embryonic β h1 gene is turned off earlier and is usually not detected after 11.5 days, that is, at the yolk sac-to-fetal liver switch in hematopoiesis (Figs. 3 and 5). Although the human γ -globin genes are the earliest to be expressed in the yolk sac, they are the only embryonic genes to remain active in the early fetal liver. They are only silenced by day 16 (Figs. 3–5) and are therefore expressed at the embryonic and fetal stage but not the adult stage of murine erythropoiesis.

The human β -globin gene was compared with the mouse globin genes and the γ genes at all stages of development. As expected from previous reports, the β -globin gene is silent at the embryonic stage (Figs. 4 and 5) but becomes active in the early fetal liver when it rises rapidly to maximum levels and remains on in the adult (Figs. 4 and 5).

Quantitation of the S1 signals and correction for the

copy number of the genes (line 15 carries two ϵ and γ genes but only one β gene) shows that the different lines express similar amounts of human globin RNA relative to the total expression of the mouse globin genes, with the exception of line 13 (Table 1). It is presently not clear why this line expresses only half of the levels observed in previous lines or the other lines in this study, but we are currently testing whether, for example, imprinting of the integration site in this mouse line is responsible for this phenomenon. Surprisingly, the main human embryonic gene (ϵ -globin) is expressed maximally at only 20–30% of the total human $\epsilon + \gamma$ level (Fig. 6), which is lower than the expression of each copy of the murine $\epsilon\gamma$ gene. The lower level of ϵ gene expression is balanced by a high level of γ -globin gene expression, which is ~70–80% of the total $\epsilon + \gamma$ signal and is higher than the murine β h1 gene. As a result, the total embryonic outputs of the mouse and the human locus are comparable. At the start of fetal liver hematopoiesis, the level of the γ gene has been reduced considerably (~15% of the total human RNA; Fig. 6) and is almost silenced by day 16 of development. The human β gene behaves very much like the mouse β genes and expresses at very similar levels per gene copy.

It is therefore clear that the human ϵ gene follows the qualitative, but not quantitative, pattern of its structural homolog $\epsilon\gamma$ in the mouse. The human β gene follows its murine homolog almost completely. The γ -globin genes are expressed throughout the embryonic period. In the early embryo, they resemble their murine homolog β h1, but are expressed at a relatively higher level. In contrast

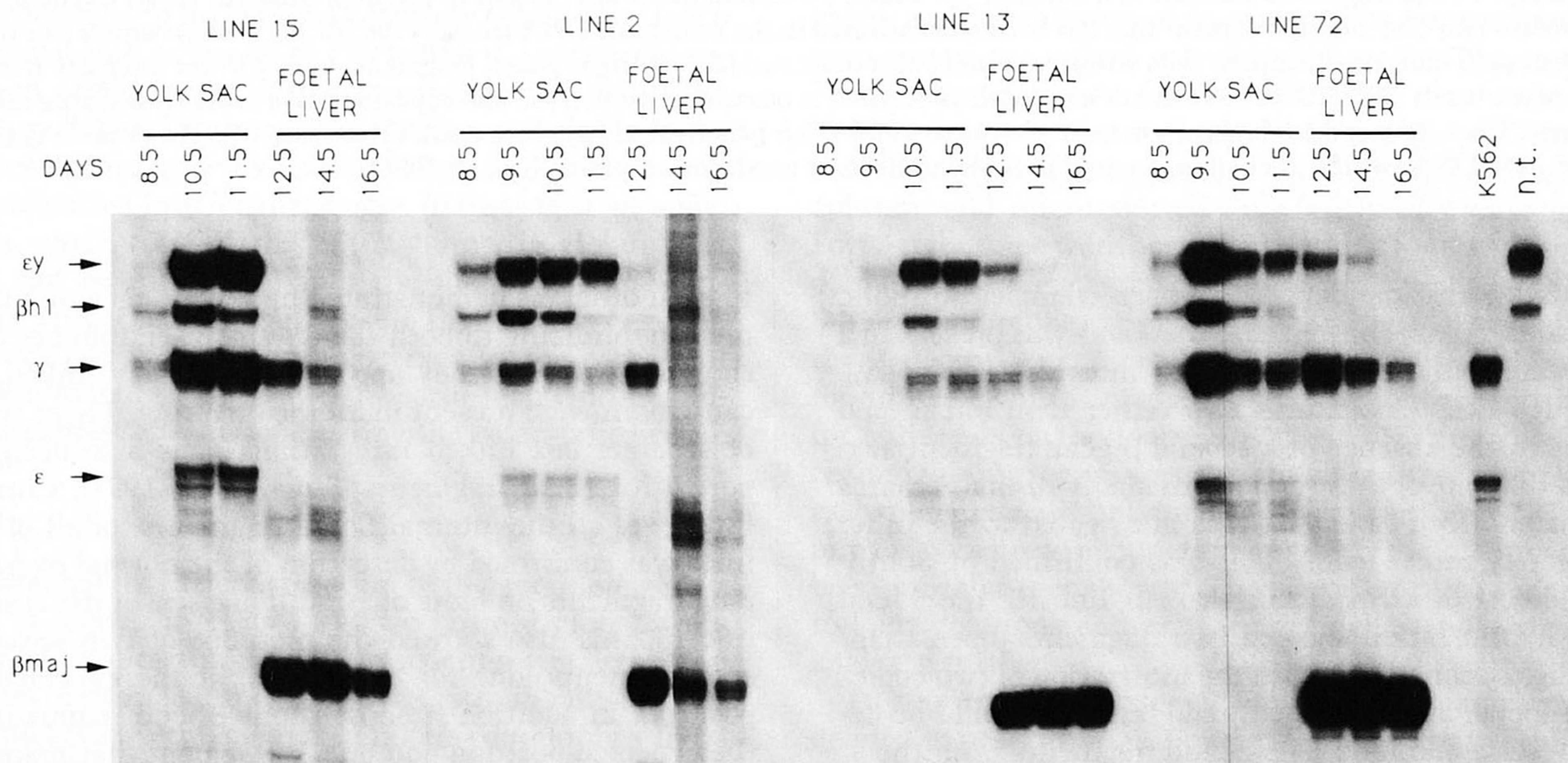


Figure 3. S1 nuclease protection analysis of the human ϵ - and γ -globin transgenes in the yolk sac and fetal liver. Each panel shows RNA isolated from transgenic embryonic yolk sac or fetal liver at different time points of development. The protected fragments are indicated at left (arrows); all probes were as described previously (Kollias et al. 1986; Fraser et al. 1990; Lindenbaum et al. 1991), with the exception of the human ϵ probe. This was a 313-bp *Bam*HI–*Bsp*MI fragment protecting a 136-nucleotide fragment from the 5' end of the ϵ RNA. Lanes 14.5 and 16.5 (days) for transgenic line 2 fetal liver contained less RNA on this gel and have been overexposed to visualize the signals. Control lanes contained RNA from K562 cells and a 10.5-day nontransgenic mouse embryo (n.t.).

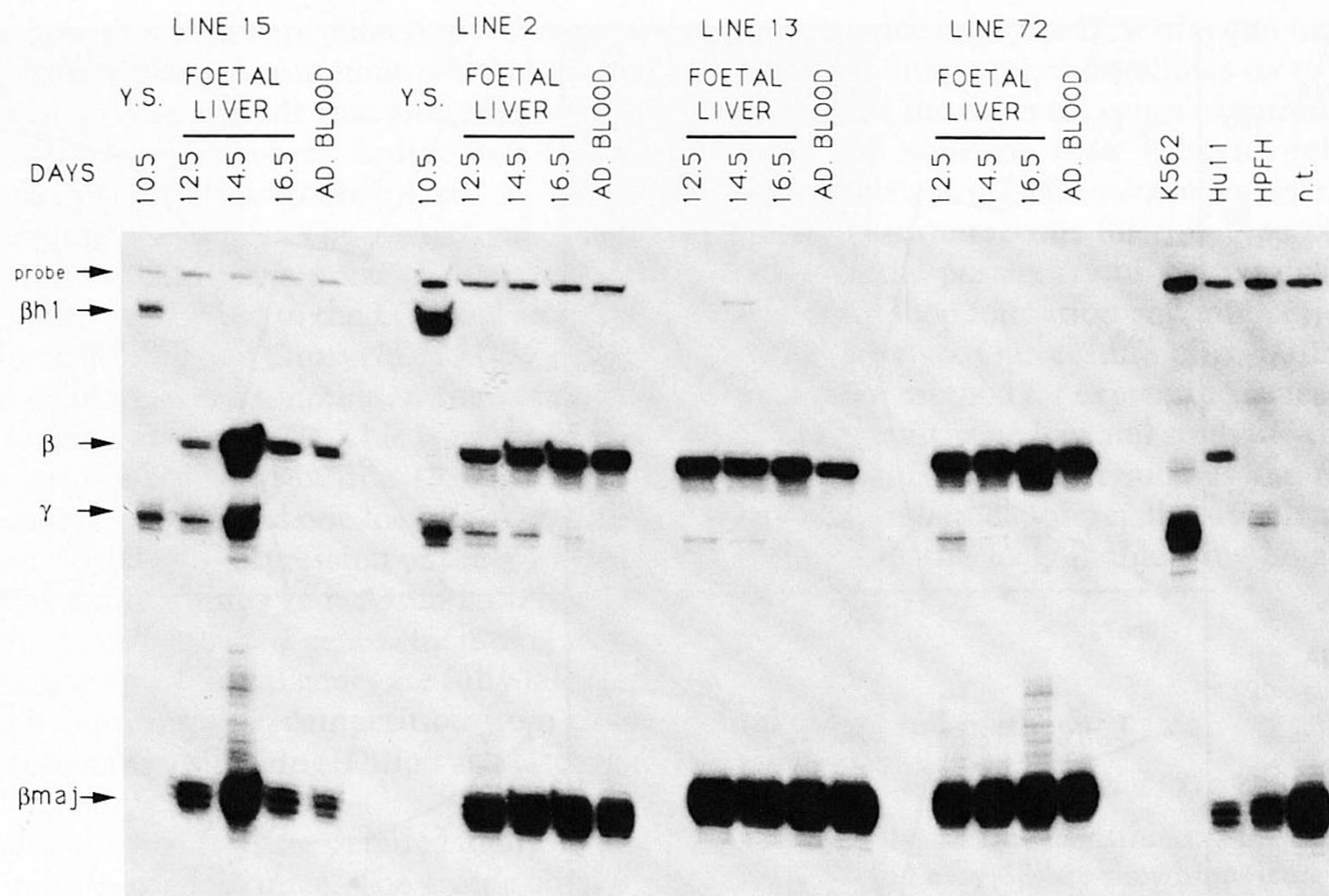


Figure 4. S1 nuclease protection analysis of the human γ and β transgenes during development. The assay was carried out as in Fig. 3. The $\epsilon\gamma$, ϵ , and 3' γ probes were replaced by a 5' γ - and β -globin probe. The reason for using a different γ probe is that not all RNAs can be detected simultaneously because of comigration of protected fragments. Control RNA samples were from K562 and Hu11 cells and from a 12.5-day nontransgenic mouse embryo (n.t.) and an HPFH (hereditary persistence of fetal hemoglobin) transgenic mouse expressing a human HPFH γ -globin gene (Berry et al. 1992).

to their murine homolog, they are switched off later and are still expressed in the early fetal liver. They are not expressed in the later fetal and early postnatal period as normally occurs in humans.

Discussion

Until now, the size limit for DNA fragments that could be used for generating transgenic mice has been governed by the maximum size of insert that could be fitted into a cosmid (~45 kb). The cosmid-joining technique we have used in this paper effectively doubles this size limit. In the present study we have generated transgenic mice carrying a 70-kb fragment, and it seems reasonable to assume that even larger fragments can be successfully injected. However, we do observe a significantly higher frequency of rearrangement of the 70-kb fragment compared with that observed for cosmid inserts. It is likely that the frequency of rearrangement (and shearing during microinjection) will increase with increasing size and ultimately determine the maximum size of fragment that can be used to generate transgenic mice at an acceptable frequency.

A further interesting feature of our results is the high frequency of transgene integration events even though the DNA was injected at very low concentrations. The large size of the injected fragment means that the number of copies injected into the nucleus is much lower than would be injected normally. For example, we would normally find it necessary to inject a 5-kb fragment at

twice the concentration used in this study to obtain the same frequency of transgenesis. This means that the number of copies injected into the nucleus in this study is ~30-fold lower, yet the frequency of integration remains similar. This suggests that the rate of integration is not directly related to the number of copies injected and that integration of large fragments is particularly efficient. However, lowering the concentration of the injected fragment appears to have lowered the rate of end-to-end ligation because the average copy number of the transgene has been reduced significantly.

The technique has allowed us to introduce a complete human β -globin locus into transgenic mice and to analyze the developmental regulation of the genes in mice in the context of the full locus. It was particularly important to determine the relative patterns of expression of the ϵ and γ genes in their normal sequence context and thereby establish whether the embryonic expression of the γ genes observed previously in mice was the result of factor differences between mice and humans or changes in sequence environment, including the absence of the ϵ gene upstream.

The most interesting result from the expression of the complete human locus in mice is the fact that the γ genes are expressed the earliest and at high levels throughout the embryonic stage. This clearly demonstrates that changes in the factor environment are at least partially responsible for their recruitment to a fetal pattern of expression in humans. However, the expression pattern of the γ genes is not identical to that ob-

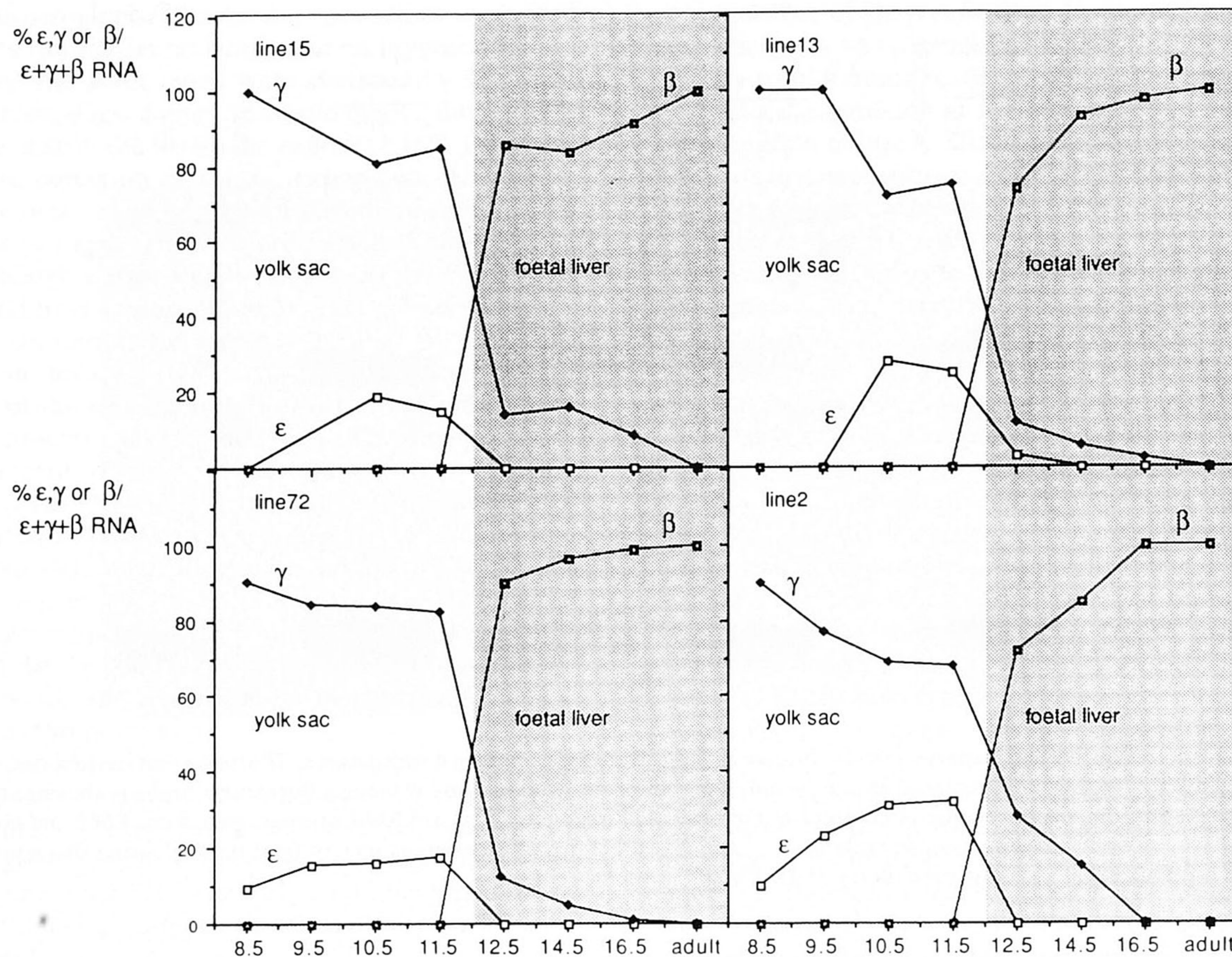


Figure 5. Expression patterns of the human ϵ -, γ -, and β -globin genes in mouse lines 15, 13, 72, and 2. The expression level of each gene was plotted as a percentage of the total human RNA at each stage of development. The levels of ϵ - and γ -globin RNA in line 15 were divided by 2 to correct for the higher copy number of these genes relative to the human β gene. The white background indicates yolk sac RNA levels; the shaded background indicates fetal liver and adult blood RNA levels.

served for either of the mouse embryonic genes. Significant levels of γ expression are consistently observed in the early fetal liver, suggesting that the regulatory sequences controlling the γ genes do not behave identically to those of the mouse embryonic genes. The fact that they are silenced later in the fetal liver demon-

Table 1. Quantitation of RNA levels observed in the different transgenic mouse lines

Time (days)	Total human RNA (% of total mouse globin RNA)			
	line 2	line 13	line 72	line 15
8.5	39	19	51	61
9.5	39	25	57	N.D.
10.5	34	36	78	80
11.5	43	41	91	81
12.5	50	30	58	62
14.5	55	26	50	47
16.5	53	28	64	60

RNA levels were determined by quantitation of the S1 signals (Figs. 3 and 4) on a Phosphor-Imager and correction for the specific activities of the probes.

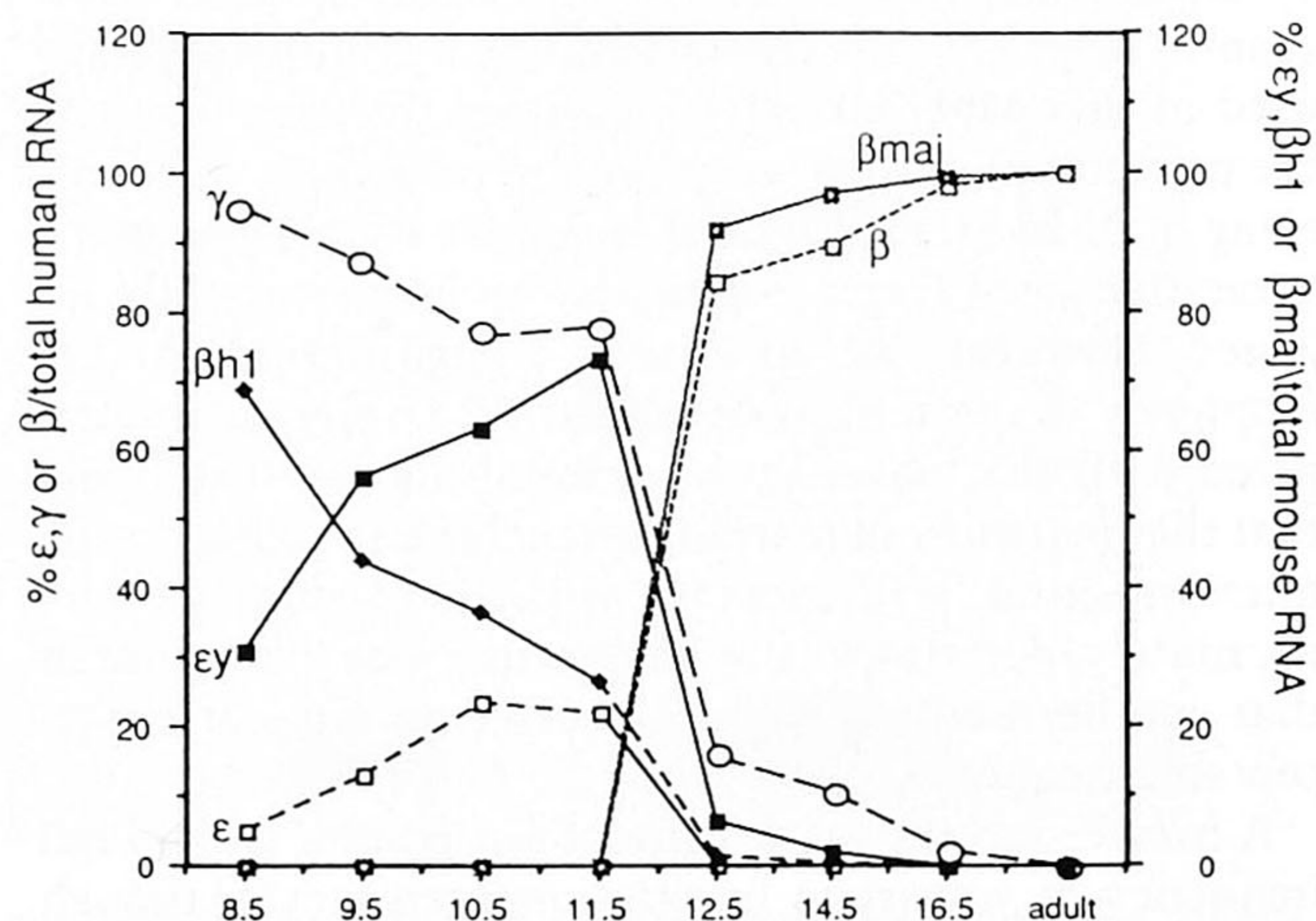


Figure 6. Expression pattern of the human and mouse globin genes. RNA levels are the average of all four mouse lines shown in Fig. 5. The RNA level for each gene was plotted for the human genes (as a fraction of total human globin RNA) and for the mouse genes (as a fraction of the total mouse globin RNA).

strates that the human γ genes are detecting a change in the balance of transcription factors that is not detected by the murine genes. We conclude that fetal recruitment of the γ genes is likely to have been caused by a combination of changes in factor environment and in the sequences surrounding the genes. The pattern of expression of the γ genes is strikingly similar to that observed when a γ gene alone was linked to the LCR and analyzed in transgenic mice (Dillon and Grosveld 1991), confirming that γ gene regulation is a function of the sequences surrounding the genes and the LCR. This is confirmed by the expression pattern observed in line 15. This mouse contains one complete locus and one locus lacking the δ and β genes, and the level of expression of the γ genes in the embryo indicates that the γ genes from both loci are expressed (Table 1). All of the γ genes, including those from the locus lacking a δ and β gene, are fully silenced in adults, which confirms that competition from these genes is not required for γ silencing (Dillon and Grosveld 1991).

This study also reports the first detailed analysis of ϵ gene expression in transgenic mice. The ϵ gene shows a pattern of expression that is qualitatively similar to that of its evolutionary homolog in the mouse, but the level of expression is much lower (20%) than the murine $\epsilon\gamma$ gene or the γ genes. On the basis of our competition model, which states that a gene proximal to the LCR should have an advantage over a gene more distal to the LCR (Dillon et al. 1991; Hanscombe et al. 1991), it could have been predicted that the ϵ gene should be expressed preferentially over the γ genes during the embryonic stage. The fact that this is not observed could be explained by several possibilities. For example, and in agreement with the model, the ϵ gene promoter could be very weak compared with that of the γ genes, and its positional advantage is not enough to provide it with higher levels of expression than the γ genes. Alternatively, there may be no competition between the ϵ and γ genes as is observed between the γ and β genes, because the genes may interact with different parts of the LCR. In such a case, the model simply does not apply and the result would be explained by the ϵ gene having a slightly weaker promoter than the γ genes in the murine environment. However, as a detailed analysis of a single ϵ gene linked to the complete LCR has not been published, it is not possible to say whether this lower level of expression is an intrinsic feature of ϵ expression in mice or whether competition from the expressed γ genes plays a role.

The expression of the human β gene follows that of β_{maj} almost completely. The only difference appears to be a slightly lower level of expression in the early fetal liver. This coincides with the fetal expression of the γ genes and suggests that this low level of γ expression leads to some competition of β expression. This is in agreement with previous observations of workers using small constructs (Hanscombe et al. 1991; Berry et al. 1992). It is not clear at present what parameters govern the level of competition in the locus. Using small constructs (Hanscombe et al. 1991), we proposed that rela-

tive distance from the LCR plays an important role. The cosmid-joining procedure allows us to test this directly by moving the different genes to various positions in the locus and studying their behavior relative to one another. Reconstruction by cosmid cassettes should be particularly advantageous for this type of experiment and provides the possibility of gaining insight into the dynamics of loop formation and interactions over long distances in vivo. Our cosmid oligo-tailing procedure also provides a method for expressing genes larger than 40 kb or other multigene loci and studying their regulation. In particular, complex loci such as the *Hox* genes, which have a multitude of regulatory elements distributed throughout the loci, should now be amenable to a detailed study.

Materials and methods

Construction of human β -globin locus

The cosmids used in reconstructing the human β -globin locus, cosLCR ϵ and cos $\gamma\gamma\delta\beta$, are shown in Figure 1 and are described in Results. Details of their construction and modifications, including the annealing and purification of the 70-kb fragment, have been submitted for publication (J. Strouboulis, N. Dillon, and F. Grosveld, in prep.).

Transgenic mice

Microinjection of the 70-kb fragment was carried out at a low concentration (0.3 ng/ μ l) into the pronuclei of fertilized mouse eggs, and transgenic mice were identified by Southern blot analysis of DNA prepared from tail biopsies (Hogan et al. 1986). Transgenic founder mice were bred, and embryos were dissected out at various time points after the detection of vaginal plugs (the day the plug was observed was taken as day 0.5). Transgenic fetuses were identified by Southern blot or slot blot analysis of placental DNA.

Mapping of the transgenic loci

Placental DNA (10 μ g) from the four transgenic lines was digested with *Eco*RI or *Hind*III and Southern blotted under alkaline conditions onto nylon filters (Zetaprobe, Bio-Rad). Human DNA as well as cosmid DNA (200 pg) diluted in nontransgenic mouse DNA were included as controls. These were hybridized to *Eco*RI- or *Hind*III-digested cosmid DNA (100–200 ng of cosLCR ϵ or cos $\gamma\gamma\delta\beta$) at 65°C. Repetitive sequences in the probes were suppressed by prehybridizing to an excess of sonicated human and mouse placental DNA, as described previously (Sealey et al. 1985). After hybridization, blots were rinsed four times in 0.3 \times SSC, 0.1% SDS, and washed in 0.1 \times SSC, 0.1% SDS, for 1 hr, all at 65°C.

RNA Analysis

RNA was extracted from embryonic yolk sacs, fetal livers, and blood and analyzed by S1 nuclease protection as described previously (Kollias et al. 1986; Fraser et al. 1990; Lindenbaum and Grosveld 1990). Quantitation of signals obtained from Southern blot and S1 nuclease protection analysis were carried out using a Phosphor-Imager (Molecular Dynamics).

Acknowledgments

We are grateful to Lucio Luzzatto for providing samples of cord blood. This work was supported by the Medical Research Council (UK).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Berry, B., F. Grosveld, and N. Dillon. 1992. A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. *Nature* **358**: 499–502.
- Collins, F.S. and S.M. Weissman. 1984. The molecular genetics of human hemoglobin. *Prog. Nucleic Acid Res. Mol. Biol.* **31**: 315–462.
- Dillon, N. and F. Grosveld. 1991. Human γ -globin genes silenced independently of other genes in the β -globin locus. *Nature* **350**: 252–254.
- Dillon, N., D. Talbot, S. Philipsen, O. Hanscombe, P. Fraser, S. Pruzina, M. Lindenbaum, and F. Grosveld. 1991. The regulation of the human β -globin locus. Gene expression and its control. In *Genome analysis* (ed. K.E. Davies and S.M. Tilghman), Vol. 2, pp. 99–118. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Driscoll, M.C., C.S. Dobkin, and B.P. Alter. 1989. γ δ β -thalassemia due to a de novo mutation deleting the 5' β -globin gene activation-region hypersensitive sites. *Proc. Natl. Acad. Sci.* **86**: 7470–7474.
- Forrester, W.C., S. Takegawa, T. Papayannopoulou, G. Stamatoyannopoulos, and M. Groudine. 1987. Evidence for a locus activator region. *Nucleic Acids Res.* **18**: 3503–3508.
- 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.
- Fraser, P., J. Hurst, P. Collis, and F. Grosveld. 1990. DNaseI hypersensitive sites 1, 2 and 3 of the human β -globin dominant control region direct position-independent expression. *Nucleic Acids Res.* **18**: 3503–3508.
- Grosveld, F., G.B. van Assendelft, D. Greaves, and G. Kollias. 1987. Position independent expression of the human beta globin gene in transgenic mice. *Cell* **51**: 975–985.
- 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.
- Hogan, B., F. Costantini, and E. Lacy. 1986. *Manipulating the mouse embryo*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Kioussis, D., E. Vanin, T. deLange, R.A. Flavell, and F.G. Grosveld. 1983. β -globin gene inactivation by DNA translocation in γ -thalassaemia. *Nature* **306**: 662–666.
- Kollias, G., N. Wrighton, J. Hurst, and F. Grosveld. 1986. Regulated expression of human A γ -, β -, and hybrid γ β -globin genes in transgenic mice: Manipulation of the developmental expression patterns. *Cell* **46**: 89–94.
- Lindenbaum, M.H. and F. Grosveld. 1990. An in vitro globin gene switching model based on differentiated embryonic stem cells. *Genes & Dev.* **4**: 2075–2085.
- Raich, N., T. Enver, B. Nakamoto, B. Josephson, T. Papayannopoulou, and G. Stamatoyannopoulos. 1990. Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* **250**: 1147–1149.
- Sealey, P., P. Whittaker, and E. Southern. 1985. Removal of repeated sequences from hybridization probes. *Nucleic Acids Res.* **13**: 1905–1922.
- Shih, D.M., R.J. Wall, and S.G. Shapiro. 1990. Developmentally regulated and erythroid-specific expression of the human embryonic β -globin gene in transgenic mice. *Nucleic Acids Res.* **18**: 5465–5472.
- Singh, G., D. Supp, C. Schreiner, J. McNeish, H. Merker, N. Copeland, N. Jenkins, S. Potter, and W. Scott. 1991. Legless insertional mutation: Morphological, molecular, and genetic characterization. *Genes & Dev.* **5**: 2245–2255.
- Stamatoyannopoulos, G. 1991. Human hemoglobin switching. *Science* **252**: 353.
- Taramelli, R., D. Kioussis, E. Vanin, K. Bartram, J. Groffen, J. Hurst, and F.G. Grosveld. 1986. $\gamma\delta\beta$ -Thalassaemias 1 and 2 are the result of a 100 kbp deletion in the human β -globin cluster. *Nucleic Acids Res.* **14**: 7017–7029.
- Townes, T. and R. Behringer. 1990. Human globin locus activation region: Role in temporal control. *Trends Genet. Sci.* **6**: 219–233.
- Tuan, D., W. Solomon, Q. Li, and I.M. London. 1985. The " β -like-globin" gene domain in human erythroid cells. *Proc. Natl. Acad. Sci.* **82**: 6384–6388.
- Wong, P., S. Chung, J. White, S. Reicheld, M. Patterson, B. Clarke, and D. Chui. 1983. Adult hemoglobins are synthesized in murine fetal hepatic erythropoietic cells. *Blood* **62**: 1280–1288.