Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 6589–6593, November 1984
Biochemistry

Triplication of a four-gene set during evolution of the goat β-globin locus produced three genes now expressed differentially during development

(gene family/embryonic and pseudoglobin genes/molecular cloning/DNA sequences)

TIM M. TOWNES, MARY C. FITZGERALD, AND JERRY B. LINGREL*

University of Cincinnati College of Medicine, Department of Microbiology and Molecular Genetics, Cincinnati, OH 45267

Communicated by Donald D. Brown, June 19, 1984

ABSTRACT **Distinct hemoglobins are synthesized in goats at different stages of development, similar to humans. Embryonic hemoglobins ($\zeta_2\varepsilon_2$ and $\alpha_2\varepsilon_2$) are synthesized initially and are followed sequentially by fetal ($\alpha_2\beta_2^F$), preadult ($\alpha_2\beta_2^C$), and adult ($\alpha_2\beta_2^A$) hemoglobins. To help understand the basis of these switches, the genes of the β-globin locus have been cloned and their linkage arrangement has been determined by the isolation of λ phage carrying overlapping inserts of genomic goat DNA. The locus extends over 120 kilobase pairs and consists of 12 genes arranged in the following order: ε^I–ε^{II}–$\psi\beta^X$–β^C–ε^{III}–ε^{IV}–$\psi\beta^Z$–β^A–ε^V–ε^{VI}–$\psi\beta^Y$–β^F. Comparison of the nucleotide sequence of the 12 genes shows that the locus is organized into three homologous four-gene sets that presumably evolved by the triplication of an ancestral set of four genes (ε–ε–$\psi\beta$–β). Interestingly, the three genes (β^C, β^A, and β^F) located at the ends of the four-gene sets are expressed at different stages of development. Therefore, the goat β^F-, β^C-, and β^A-globin genes appear to have evolved by a mechanism that includes the triplication of 40–50 kilobase pairs of DNA and the recruitment of newly formed genes for expression in fetal, preadult, and adult life.**

Goats, like humans, switch hemoglobins during development. Both synthesize embryonic hemoglobins composed of either ζ or α and ε polypeptides ($\zeta_2\varepsilon_2$ or $\alpha_2\varepsilon_2$) in the yolk sac blood islets. The embryonic globin is then replaced by fetal hemoglobin ($\alpha_2\beta_2^F$ in goats and $\alpha_2\gamma_2$ in humans) when the major site of erythropoiesis switches to the fetal liver (1–3). The non-α polypeptide of the goat fetal hemoglobin will be referred to as β^F rather than the previously used γ (4) to reflect its adult β-like origin as indicated by both amino acid and nucleotide sequence data (4–6) and by its evolutionary history, which is described in this paper. Although humans switch directly to adult hemoglobin ($\alpha_2\beta_2$) at birth, goats undergo an additional switch to a juvenile or preadult hemoglobin composed of two α and two β^C polypeptides ($\alpha_2\beta_2^C$). This hemoglobin is then replaced by adult hemoglobin ($\alpha_2\beta_2^A$) at approximately 3 months of age. Interestingly, this final switch is reversible; complete replacement of adult hemoglobin by preadult hemoglobin can be induced in adult goats by anemic stress, hypoxia, or injection of erythropoietin (7–9).

The isolation and characterization of 10 β-like globin genes of the goat has been reported (10–16). Nine of these genes, in the order ε^I–ε^{II}–$\psi\beta^X$–β^C–ε^{III}–ε^{IV}–$\psi\beta^Z$–β^A–ε^V, were isolated as overlapping inserts in Charon 4A clones (16). Comparison of the nucleotide sequences of the first 8 genes in this cluster showed that these genes are organized into two four-gene sets presumably formed by the duplication of an ancestral four-gene set (ε–ε–$\psi\beta$–β). Each gene in one set is most homologous to the gene occupying the same position in the other four-gene set. Thus, ε^I, and ε^{III}, ε^{II} and ε^{IV}, $\psi\beta^X$ and $\psi\beta^Z$, and β^C and β^A are >90% homologous, both in coding and noncoding sequence. Several observations suggested that a third four-gene set might be present in the goat β-globin locus. First, the ε^V gene is linked downstream of β^A and is highly homologous to ε^I and ε^{III} (16). Therefore, ε^V could begin a final set of four genes. Second, the fetal globin gene (β^F) is highly homologous to β^C and β^A (6) and therefore could represent the fourth gene in the last four-gene set. Finally, two globin-hybridizing sequences that were not represented in the Charon 4A clones have been observed in genomic Southern blots (16). One of the sequences hybridized to a pseudogene-specific probe. Therefore, these two sequences could represent an embryonic gene and a pseudoglobin gene that would complete the final four-gene set. In this paper, we report the isolation of genomic clones that contain an embryonic (ε^{VI}) gene and a pseudoglobin ($\psi\beta^X$) gene and link the fetal globin gene with the rest of the β-globin cluster.

EXPERIMENTAL PROCEDURES

A partial *Sau*3A library of genomic goat DNA was constructed in the λ vector EMBL 4 (17) by standard procedures (18). Without amplification, 600,000 recombinant phage were screened in duplicate as described by Benton and Davis (19) with a mixture of ε^I-5′, $\psi\beta^Z$-5′, and β^F-5′ probes (16). Appropriate fragments of isolated phage clones were subcloned into the plasmid pUC8 (20) and sequenced by the method of Maxam and Gilbert (21).

RESULTS

Linkage Map of the Goat β-Globin Gene Family. To isolate clones encompassing the entire β-globin locus, we constructed a partial *Sau*3A library of genomic goat DNA in the λ vector EMBL 4 (17) and screened it with a mixture of embryonic, pseudo-, and fetal globin probes (16). Twenty-one different globin-hybridizing clones were obtained. Four of these clones (nos. 31, 32, 33, and 34) are illustrated in Fig. 1, along with the previously identified (16) Charon 4A clones. Clone 32 contains two globin-hybridizing *Eco*RI fragments [12 and 2.3 kilobase pairs (kb)], which correspond to globin sequences previously observed in genomic Southern blots (16) but not represented in the Charon 4A clones. The 2.3-kb fragment of clone 32 hybridizes specifically with the pseudogene probe (data not shown). Clone 34 contains this same 2.3-kb *Eco*RI fragment linked upstream from the fetal globin gene. Therefore, the four newer clones illustrated in Fig. 1 contain two previously unreported globin-hybridizing sequences and complete the linkage map of the goat β-globin gene family.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: kb, kilobase pair(s).

*To whom reprint requests should be addressed.

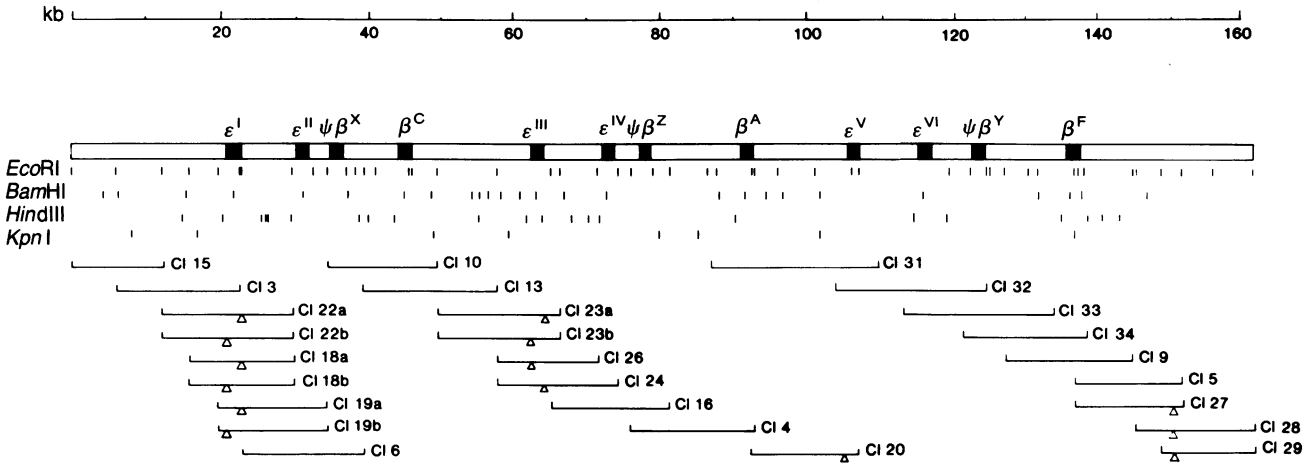

FIG. 1. Linkage map of the goat β-globin locus. Clones 31–34 were isolated from a partial *Sau*3A library constructed in the λ vector EMBL 4 (17). All of the other clones were isolated from a partial *Eco*RI library constructed in Charon 4A as described (10). The triangles represent sites of insertion of the πVx plasmid in Charon 4A clones isolated prevously by πVx recombinant screening (15, 16). Overlaps between phage clones were confirmed by digesting the purified overlapping *Eco*RI fragments with *Hae* III, Rsa I, or both and comparing the resulting fragments by electrophoresis on agarose gels. All of the genes are oriented in the same direction, left to right, on the map.

Identification of a Sixth Embryonic Sequence. Part of the nucleotide sequence of the gene in the 12-kb *Eco*RI fragment of clone 32 is compared with the sequence from −122 to +139 of ε^{II} and ε^{IV} in Fig. 2*a*. The first exon of the newly isolated gene is ≈95% homologous to the embryonic genes that occupy the same relative positions in the β^{C} (ε^{I}, ε^{II}, $\psi\beta^{X}$, β^{C}) and β^{A} (ε^{III}, ε^{IV}, $\psi\beta^{Z}$, β^{A}) four-gene sets. This high degree of sequence similarity in the first exon establishes the embryonic nature of the newly isolated gene, which will subsequently be referred to as ε^{VI}. Since the first exon of ε^{VI} is ≈95% homologous to the first exons of ε^{II} and ε^{IV} but only 80% homologous to those of ε^{I}, ε^{III}, and ε^{V}, the newly isolat-

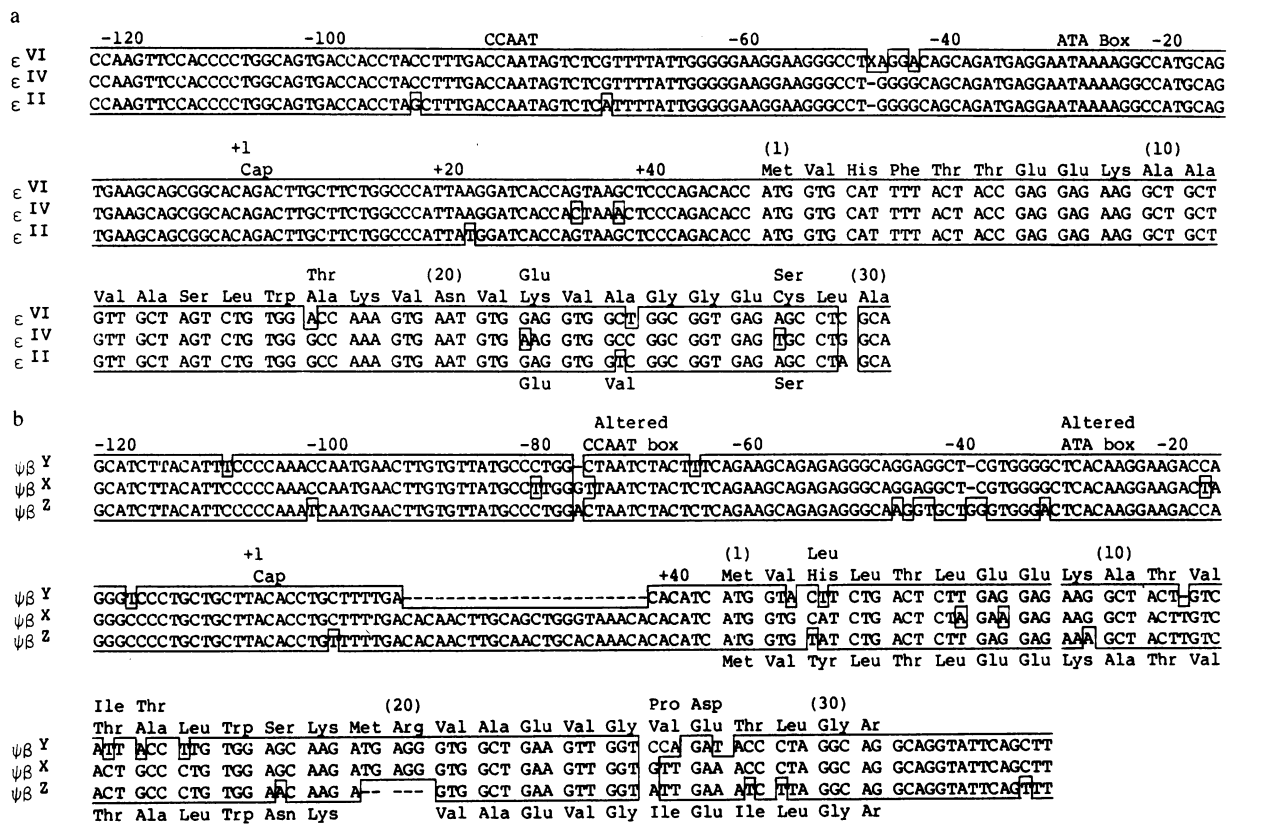

FIG. 2. Comparison of the nucleotide sequence of the goat ε^{II}, ε^{IV}, and ε^{VI} (*a*) and $\psi\beta^{X}$, $\psi\beta^{Y}$, and $\psi\beta^{Z}$ (*b*) genes from positions −122 to +139 and −122 to +150, respectively. The 6.0- and 2.3-kb *Eco*RI fragments of clone 33 and 34, respectively, were subcloned into pUC8 (20). The 6.0-kb fragment is actually a partial *Sau*3A/*Eco*RI genomic fragment; the *Eco*RI site immediately adjacent to the *Sau*3A site is derived from the vector arm. The sequence of ε^{VI} was determined in both directions from the 5′-end-labeled *Mst*EII site at −47. The sequence of $\psi\beta^{Y}$ was determined in both directions from the 5′-end-labeled *Pvu* II site at +181. Regions of homology between the genes are boxed, and the deduced amino acid sequence is indicated above the exon for ε^{VI}, ε^{IV}, $\psi\beta^{Y}$, and $\psi\beta^{X}$ and below the exon for ε^{II} and $\psi\beta^{Z}$. Dashes represent deletions inserted to maximize homology and the single X represents an unreadable base. The sequence of ε^{II} is from Shapiro *et al.* (15); ε^{IV} is from Townes *et al.* (16). $\psi\beta^{X}$ and $\psi\beta^{Z}$ sequences are from Cleary *et al.* (14).

ed gene appears to be closely related to the second-position genes in the β^C and β^A four-gene sets. The similarity of the three second-position genes is even more evident when the noncoding sequences are compared with each other and with the other genes in the locus. The 5′ untranslated region and the first 120 base pairs of 5′ flanking sequence of ε^{II}, ε^{IV}, and ε^{VI} are >95% homologous and are <50% homologous to the other genes in the cluster.

Identification of a Third Pseudoglobin Gene. Part of the nucleotide sequence of the gene contained in the 2.3-kb *Eco*RI fragment of clone 34 is compared with the sequence from −122 to +150 of $\psi\beta^X$ and $\psi\beta^Z$ in Fig. 2*b*. This gene is >90% homologous to $\psi\beta^X$ and $\psi\beta^Z$, both in coding and in noncoding regions, and therefore appears to be closely related to the two pseudogenes. Moreover, it contains some of the same defects that make $\psi\beta^X$ and $\psi\beta^Z$ nonfunctional. For instance, the newly isolated sequence possesses the same alterations in the "ATA" box as $\psi\beta^X$ and $\psi\beta^Z$. The ATA box at −35 in the functional goat globin genes has the sequence A-A-T-A-A-A-A. This sequence is mutated to C-T-C-A-C-A-A in $\psi\beta^X$, $\psi\beta^Z$, and the newly isolated gene. Since the newly isolated gene is highly homologous to $\psi\beta^X$ and $\psi\beta^Z$ and shares with these two genes an identically mutated ATA box, it will subsequently be referred to as $\psi\beta^Y$. $\psi\beta^Y$ also shares with $\psi\beta^X$ and $\psi\beta^Z$ the same splice sequence defect in the donor site of the first intron. All three genes contain a G-C at this splice junction as opposed to the G-T dinucleotide found in most mammalian genes transcribed by RNA polymerase II (22), including all of the other goat β-globin genes. $\psi\beta^Y$, like $\psi\beta^X$ and $\psi\beta^Z$, also contains a mutated C-C-A-A-T box. All actively functioning globin genes reported to date contain the sequence C-C-A-A-T at a position approximately 80 base pairs upstream from the transcriptional start site (23–25). This sequence, which is essential for optimal promoter efficiency (26), is mutated to C-T-A-A-T in $\psi\beta^Y$ and $\psi\beta^Z$ and to T-T-A-A-T in $\psi\beta^X$. Interestingly, $\psi\beta^Y$ does not share with $\psi\beta^X$ and $\psi\beta^Z$ in the single-base (thymidine) insertion between codons 11 and 12. This insertion in $\psi\beta^X$ and $\psi\beta^Z$ produces a frameshift mutation that results in the formation of a stop codon in the first exon.

The fact that $\psi\beta^Y$ is highly homologous to $\psi\beta^X$ and $\psi\beta^Z$ and shares with these latter genes identical point mutations strongly suggests that $\psi\beta^X$, $\psi\beta^Y$, and $\psi\beta^Z$ were produced by a recent triplication of a preexisting pseudoglobin gene.

DISCUSSION

Evolution of a Gene Family. The data presented above illustrate the linkage of 12 genes in the β-globin locus of the goat genome and strongly suggest that the locus was formed relatively recently by the triplication of 40–50 kb of DNA containing a basic four-gene set (ε–ε–$\psi\beta$–β). A model for the evolution of the goat β-globin locus is shown in Fig. 3. After the formation of an ancestral four-gene set composed of two embryonic and two β-globin genes, the first adult gene became defective, presumably due to mutations in the C-C-A-A-T or ATA box. This basic four-gene set was then duplicated to form a second set of four genes. Interestingly, the gene occupying the fourth position in the new four-gene set was recruited for expression during fetal development and is now designated β^F. After the duplication that formed the β^F four-gene set, a third four-gene set was formed, probably as the result of an unequal crossing-over event. One possible site of recombination is illustrated in Fig. 3. Subsequent to the crossover, the developmental specificity of the fourth position gene in the first four-gene set was changed. This gene (β^C) is now expressed during preadult development.

Several observations suggest that the β^F four-gene set was formed before the β^C set. First, comparison of the nucleotide sequence of the β^C, β^A, and β^F genes shows that, although all three of these genes are highly homologous, β^F has diverged more than β^A and β^C (4). Second, the single-base-pair insertion present in $\psi\beta^X$ and $\psi\beta^Z$ but not in $\psi\beta^Y$ implies that the four-gene set containing $\psi\beta^Y$ was formed prior to the insertion. That is, the extra base was inserted into the $\psi\beta^X/\beta^Z$ precursor after the initial duplication. This gene was then duplicated in the unequal crossing-over event that produced the third four-gene set. Although the precise site of the recombination cannot be determined without additional sequence data, the presence of the inserted thymidine in both $\psi\beta^X$ and $\psi\beta^Z$ is consistent with a cross-over in a region of DNA 5′ to the $\psi\beta^X/\psi\beta^Z$ precursor.

Three observations suggest that the triplicated locus is most likely present in the entire caprine population. The same *Eco*RI pattern of globin-hybridizing sequences has been detected in genomic Southern blots of goats obtained from at least five different herds. Second, the two genomic libraries from which the goat globin genes were isolated were prepared from goats obtained from different herds. The

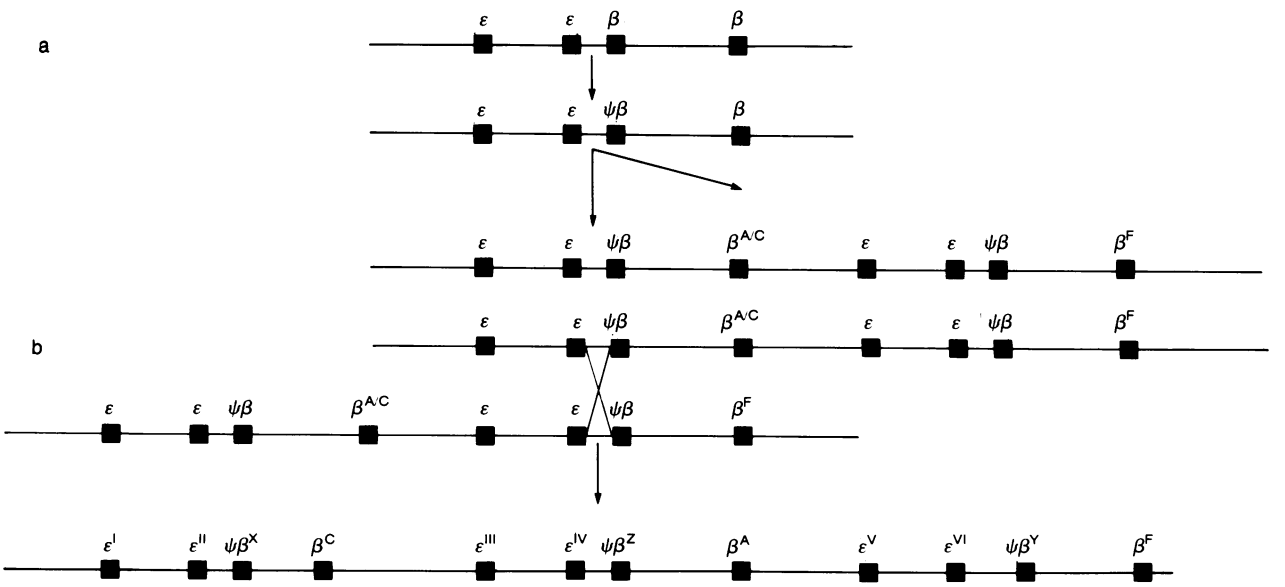

FIG. 3. Model for the evolutionary history of the goat β-globin locus. (*a*) Formation of the ancestral pseudogene and subsequent duplication of the four-gene set. (*b*) Illustration of the unequal crossing-over event that most likely produced the triplicated locus.

Charon 4A library contained 10 of the 12 genes and the EMBL 4 library contained almost the entire β-globin locus, lacking only 5 kb of DNA around the ε^{III} gene. Finally, the triplication most likely preceded the divergence of sheep and goats. Sheep, like goats, possess both fetal and preadult globin polypeptides that are adult β-like in sequence (4, 5, 27). Therefore, the sheep fetal, preadult, and adult β-globin genes appear to have evolved by the triplication of an adult β-like sequence. Although the sheep preadult gene has not been isolated, the nucleotide sequences of the sheep fetal and adult β-globin genes have been determined (28). These genes are highly homologous to the goat fetal and adult genes in both coding and noncoding sequences. Furthermore, the amino acid sequences of preadult (β^C) globin gene products are similar in sheep and goats. Therefore, the triplication event that produced the β^C, β^A, and β^F genes most likely occurred in an ancestor common to sheep and goats. This is in agreement with the evolutionary tree derived by Czelusniak *et al.* (29) using the method of maximum parsimony.

Comparison of β-Globin Loci. The structures of four mammalian β-globin loci (30–34) are compared in Fig. 4. Of all the β-globin gene families, the goat family is by far the largest. As discussed above, the goat locus contains 12 linked genes that span >120 kb of DNA. Like other mammalian β-globin gene families, all of the goat β-globin genes are oriented in the same direction; left to right on the linkage map. However, unlike the human β-globin family, the goat genes are not organized in the order in which they are expressed during development. The adult gene β^A is positioned in the middle of the cluster with the preadult (β^C) and fetal (β^F) genes located on either side at ≈50-kb intervals. This organization precludes a mechanism of gene activation that relies on sequential opening of adjacent chromosomal domains throughout the locus during development.

Primordial Four-Gene Set. Shapiro *et al.* (15) have shown that an individual four-gene set in the goat β-globin locus is similar to the β-globin loci of several other mammals and suggested that a primordial four-gene set may have existed before the mammalian radiation. Hardison (34) has also made this suggestion afer comparing the positions and sequences of rabbit and human β-globin genes. The first-position gene in an individual four-gene set in the goat locus is highly homologous to human ε (35), rabbit β4 (34), and mouse εy3 (36). Similarly, the fourth-position gene is highly homologous to human β (32), rabbit β1 (31), and the duplicated mouse adult genes β^{maj} and β^{min} (37). Although the origins of the second- and third-position genes in the goat locus are somewhat obscured by differences in the rate at which first- and second-position genes fix mutations (15) and by gene conversions that homogenize third- and fourth-position genes (38), ε^{II}, ε^{IV}, and ε^{VI} appear to be related to human γ and rabbit β3 (15, 39) and goat $\psi\beta^X$, $\psi\beta^Y$, and $\psi\beta^Z$ are most likely related to human δ and rabbit ψβ2 (38).

Convergent Evolution of Human and Goat Fetal Globin Genes. As indicated above, the genes that now encode the goat preadult and fetal globin polypeptides were formed relatively recently by two separate duplications of an adult β-like precursor. The adult β-like origin of the goat fetal globin gene is very different from the origin of the human fetal globin gene, which presumably originated from an embryonic globin precursor (34, 39). Therefore, these two fetal globin genes have distinct evolutionary histories. Both genes encode polypeptides that possess higher oxygen affinities than their adult counterparts (1). Thus, both fetal proteins are able to efficiently accept oxygen from the maternal adult hemoglobin. However, the two proteins have very different amino acid sequences. The distinct origin of the two fetal globin genes therefore seems to represent a clear example of convergent evolution.

Developmental Regulation of β^C, β^A, and β^F. Perhaps the most interesting aspect of the two duplications that formed the goat β-globin locus is that these duplications mark the origins of the fetal and preadult globin genes. It is possible that the duplications moved these sequences into chromosomal domains that are opened during fetal or preadult life. In this case, the newly formed adult-like genes would be expressed during fetal or preadult life as a result of their new locations in the genome. Alternatively, mutations in regulatory sequences may have changed the developmental specificity of these genes subsequent to the duplications that formed them. In this respect, it is interesting that the nucleotide sequences of the β^C and β^A genes are identical from −79 to +56 (11). The first nucleotide differences at the 5′ ends of these genes is at position −80 and there are only eight other nucleotide differences through position −227. The fetal globin gene is also highly homologous to β^A in the 5′ flanking region. These two genes differ by only three nucleotides between −79 and +56 and there are only 10 other differences through −227. In fact, all three genes are >90% homologous as far as 700 nucleotides upstream from their transcriptional start sites (40). Therefore, if developmental regulatory sequences are located at the 5′ ends of these genes, single-base-pair differences must be sufficient for differential control.

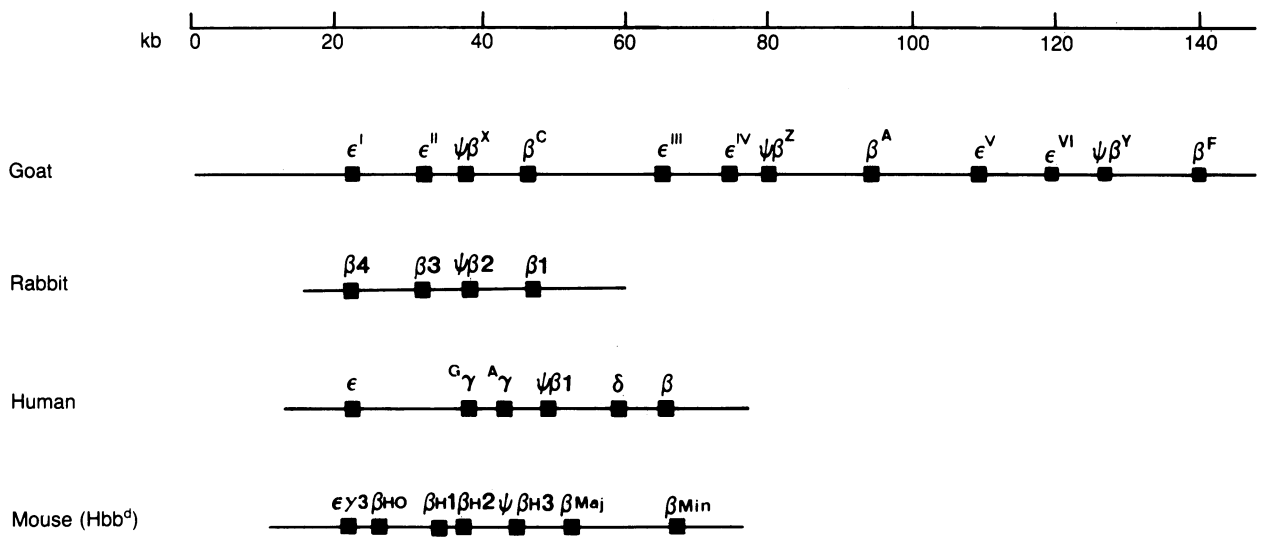

FIG. 4. Comparison of four mammalian β-globin loci.

We wish to acknowledge the excellent technical assistance of Ms. Jane Selegue. This work was supported by Grants HL 15996 and GM 10999 from the National Institutes of Health.

1. Kitchen, H. & Brett, I. (1974) *Ann. N.Y. Acad. Sci.* **241,** 653–671.
2. Huisman, T. H. J. (1974) *Ann. N.Y. Acad. Sci.* **241,** 549–555.
3. Weatherall, D. S. & Clegg, J. B. (1979) *Cell* **16,** 467–479.
4. Schon, E. A., Cleary, M. L., Haynes, J. R. & Lingrel, J. B. (1981) *Cell* **27,** 359–369.
5. Wilson, J. B., Adams, H. R. & Huisman, T. H. J. (1969) *Biochim. Biophys. Acta.* **181,** 367–372.
6. Dayhoff, M. O. (1972) *Protein Sequence Database: Atlas of Protein Sequence and Structure*: Version 5. (National Biomedical Research Foundation, Silver Spring, MD), p. 27, D-371/D372.
7. Huisman, T. H. J., Lewis, J. P., Blunt, M. H., Adams, H. R., Miller, A., Dozy, A. M. & Boyd, E. M. (1969) *Pediatr. Res.* **3,** 189–198.
8. Thurman, T. F., Boyer, S. H., Crosby, E. F., Shepard, M. K., Noyes, A. N. & Stohlman, F., Jr. (1970) *Blood* **36,** 598–606.
9. Tucker, E. M. (1971) *Biol. Rev.* **46,** 341–386.
10. Robbins, J., Rosteck, P., Jr., Haynes, J. R., Freyer, G., Cleary, M. L., Kalter, H. D., Smith, K. & Lingrel, J. B. (1979) *J. Biol. Chem.* **254,** 6187–6195.
11. Haynes, J. R., Rosteck, P., Jr., Schon, E. A., Gallagher, P. M., Burke, D. J., Smith, K. & Lingrel, J. B. (1980) *J. Biol. Chem.* **255,** 6355–6367.
12. Haynes, J. R., Rosteck, P., Jr., & Lingrel, J. B. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 7127–7131.
13. Cleary, M. L., Haynes, J. R., Schon, E. A. & Lingrel, J. B. (1980) *Nucleic Acids Res.* **8,** 4791–4802.
14. Cleary, M. L., Schon, E. A. & Lingrel, J. B. (1981) *Cell* **26,** 181–190.
15. Shapiro, S. G., Schon, E. A., Townes, T. M. & Lingrel, J. B. (1983) *J. Mol. Biol.* **169,** 31–52.
16. Townes, T. M., Shapiro, S. G., Wernke, S. M. & Lingrel, J. B. (1984) *J. Biol. Chem.* **259,** 1896–1900.
17. Murray, N. E. (1983) in *Lambda II*, eds. Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 395–432.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 270–294.
19. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
20. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
21. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
22. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
23. Lacy, E. & Maniatis, T. (1980) *Cell* **21,** 545–553.
24. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21,** 653–668.
25. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8,** 127–142.
26. Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. & Weissmann, C. (1983) *Cell* **32,** 695–706.
27. Huisman, T. H. J., Adams, H. R., Dimmock, M. D., Edwards, W. E. & Wilson, J. B. (1967) *J. Biol. Chem.* **242,** 2534–2541.
28. Kretschmer, P. J., Coon, H. C., Davis, A., Harrison, M. & Nienhuis, A. (1981) *J. Biol. Chem.* **256,** 1975–1982.
29. Czelusniak, J., Goodman, M., Hewett-Emmett, D., Weiss, M. L., Venta, P. J. & Tashian, R. E. (1982) *Nature (London)* **298,** 297–300.
30. Lacy, E., Hardison, R. C., Quon, D. & Maniatis, T. (1979) *Cell* **18,** 1273–1283.
31. Hardison, R. C., Butler, E. T., III, Lacy, E., Maniatis, T., Rosenthal, N. & Efstratiadis, A. (1979) *Cell* **18,** 1285–1297.
32. Fritsch, E. F., Lawn, R. M. & Maniatis, T. (1980) *Cell* **19,** 959–972.
33. Jahn, C. L., Hutchinson, C. A., Phillips, S. J., Weaver, S., Haigwood, N. L., Voliva, C. F. & Edgell, M. H. (1980) *Cell* **21,** 159–168.
34. Hardison, R. C. (1983) *J. Biol. Chem.* **258,** 8739–8744.
35. Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21,** 621–626.
36. Hansen, J. N., Konkel, D. A. & Leder, P. (1982) *J. Biol. Chem.* **257,** 1048–1052.
37. Konkel, D. A., Maizel, J. V., Jr., & Leder, P. (1979) *Cell* **18,** 1273–1283.
38. Hardies, S. C., Edgell, M. H. & Hutchinson, C. A. (1984) *J. Biol. Chem.* **259,** 3748–3756.
39. Hardison, R. C. (1981) *J. Biol. Chem.* **256,** 11780–11786.
40. Lingrel, J. B., Townes, T. M., Shapiro, S. G., Spence, S. E., Liberator, P. A. & Wernke, S. M. (1983) in *Globin Gene Expression and Hematopoietic Differentiation*, eds. Stamatoyannopolous, G. & Nienhuis, A. (Liss, New York), pp. 131–139.