MOLECULAR AND CELLULAR BIOLOGY, May 1985, p. 1025–1033
0270-7306/85/051025-09$02.00/0

Vol. 5, No. 5

Characterization, Expression, and Evolution of the Mouse Embryonic ζ-Globin Gene

AYA LEDER,* LAWRENCE WEIR, AND PHILIP LEDER

Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115

Received 19 October 1984/Accepted 15 February 1985

We have determined the complete sequence of the embryonic α-like, ζ (zeta)-globin gene of the BALB/c mouse. The structure of this gene establishes the amino acid sequence of the mouse embryonic ζ-globin polypeptide chain and allows us to identify sequences within the gene that may be important for its expression. One of these is a 300-base segment that is tightly conserved between mice and humans and is located at the 5′ end of the ζ-globin gene. By introducing the cloned gene into permanently transfected mouse erythroleukemic cell lines and comparing its transcript with that of ζ-globin mRNA derived from embryonic yolk sac erythrocytes, we are able to show that the cloned gene is transcriptionally active and that its transcript is correctly initiated and processed. Interestingly, the ζ-globin gene is also active when permanently transfected into an immunoglobulin-producing B-cell, a cell that presumably has tissue-specific requirements for gene expression. Further, a comparison of the amino acid coding sequence of the mouse ζ-globin gene to that of ζ-like globin genes of other species supports a revised evolutionary lineage in which goats and humans are closely related, whereas mice are further removed.

The α-globin locus in the mouse consists of three functional genes spread along 25 kilobases (kb) of DNA on chromosome 11 (23). The most 5′ of these, the ζ gene, is an embryonic gene that is expressed only in the nucleated erythrocytes of the yolk sac of the developing embryo. Its expression begins at day 8 of gestation, peaks at day 10, and is almost undetectable by day 15. At ca. day 12 of gestation, the nucleated yolk cells also begin to express the adult α-globin gene. In addition, the site of hematopoiesis switches at about this time from the yolk sac to the fetal liver in which only the adult genes are active. The fetal liver continues to produce nonnucleated erythrocytes containing adult α-globin polypeptides until birth, after which the bone marrow gradually becomes the major hematopoietic organ (2, 6, 9, 10).

As a first step in understanding the molecular basis for the temporal expression of these genes, a physical map of the locus was established (23) as was a detailed structure of the two adult genes, α_1 (30) and α_2 (A. Leder, manuscript in preparation). The results of these investigations established that the adult α-globin genes are encoded in three segments of DNA separated by two small intervening sequences. Similar organizational features have been found in all adult α-like globin genes examined thus far (13, 14, 21, 26, 28).

In this study, we establish the complete sequence of the remaining embryonic α-like gene, a gene we now call mouse ζ because of its evident homology to the human embryonic α-like gene, human ζ (33). These genes are quite different from the adult α genes. In addition to inferring the amino acid sequence of the mouse ζ polypeptide, we are able to show that the cloned mouse ζ gene is actively expressed after transfection into mouse erythroleukemic and B-cells. The active mouse gene retains extensive homology to the human gene in two regions, one corresponding to the coding sequences (and doubtless retained by selection for coding function) and another 5′ to the coding region. The latter is interesting because of its potential function as a regulatory region. Further sequence comparisons by using the technique of parsimony analysis allow us to suggest a small revision of the evolutionary lineage of mice, humans, and goats, in which goats and humans are closer to one another than either is to mice.

MATERIALS AND METHODS

Cloning. The mouse ζ-globin gene was cloned from an *Eco*RI digest of DNA derived from BALB/c mice with the bacteriophage vector λgtWES (24). The recombinant phage containing a 3.8-kb *Eco*RI fragment encoding the mouse ζ gene was detected with a probe derived from the human ζ-globin cDNA clone (kindly provided by Bernard Forget) and was packaged, grown, and purified by using established techniques (36). After hybridization, filters were washed at 40°C in 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate).

Sequencing. Sequencing was carried out by using both the chemical degradation technique of Maxam and Gilbert (27) and the M13 dideoxynucleotide termination method essentially as described in the *Bethesda Research Laboratories Users Manual*. Sequencing was carried out on both strands over most of the sequence; however, when opposite strand data was not obtained, the relevant fragment was sequenced twice.

Heteroduplex analysis. The mouse and human (33) genomic clones were subcloned as *Eco*RI and *Hin*dIII fragments, respectively, into the *Eco*RI and *Hin*dIII sites of pBR322 in the same orientation. They were both linearized with *Sal*I (at a unique site in pBR322) and then denatured with alkali at room temperature. Renaturation took place in 50% formamide–12 mM EDTA–50 mM NaCl–10 mM Tris (pH 8.5) at 37°C for 1 h.

Transfection. Transfection was carried out by the electroporation technique described previously (31, 37). Exponentially growing cells were centrifuged, washed, and suspended at a concentration of 1×10^7 to 2×10^7 cells per ml. Linearized transfecting DNA was added to a final concentration of 10 μg/ml (i.e., a total of 5 μg). A high-voltage pulse was applied to the mixture of cells and DNA which was then allowed to sit for 10 min at 0°C before the addition of growth

* Corresponding author.

medium. Selection for transfectants was by the pSV-gpt system (based upon the *E. coli* enzyme, guanosine phosphoribosyltransferase) described by Mulligan and Berg (29). After 48 h, the cells were selected in medium supplemented with 1 μg of mycophenolic acid per ml, the compound to which the transfected DNA conferred resistance.

Transfecting DNA. pSV7gpt* is pBR327 with the *Escherichia coli gpt** gene inserted at the *Bam*HI site (*gpt** is a modification by Mitch Raff and Martin Rosenberg of the original guanosine phosphoribosyltransferase gene). The 3.8-kb genomic *Eco*RI fragment containing the ζ-globin gene was inserted into the *Eco*RI site of pSV7gpt*. To increase the transfection efficiency, plasmid DNA was linearized with the restriction enzyme *Cla*I (31).

Cells and tissue. Two mouse erythroleukemia cell lines were used: GM-979 and SI (super inducer). M12 is an immunoglobulin-producing B-cell line established from a spontaneous B-cell lymphoma occurring in a BALB/c mouse (20). The yolk sac was removed from 10- and 12-day-old embryos of CD-1 mice and rinsed in cold Hanks buffered saline.

Induction conditions. Cells were grown in medium supplemented with butyric acid at 1.2 mM for 4 days or dimethyl sulfoxide (DMSO) at 1.5% for 4 days.

RNA isolation. Cell pellets were directly vortexed with 4 M guanidine thiocyanate–2.5 mM sodium citrate (pH 7)–0.1 M 2-mercaptoethanol and passed several times through a no. 26 0.5-in sterile needle. In both cases, total RNA was isolated by the procedure of Chirgwin et al. (8) with a CsCl gradient modification.

S1 nuclease. Single-stranded, uniformly labeled DNA probes were prepared as described by Ley et al. (25). Primer extended M-13 clones were restricted with appropriate enzymes to yield a desired fragment of the cloned insert. Labeled single-stranded fragments were then isolated on preparative alkaline agarose minigels (1% agarose in 0.03 M NaOH, 2 mM EDTA). The electroeluted probe was hybridized to total RNA with a modified procedure of Berk and Sharp (5). The hybridization mixture contained 50,000 to 100,000 cpm of probe (SA = 10^8 cpm/μg; 10 μg of total cellular RNA, 75% formamide, 400 mM NaCl, 20 mM Tris [pH 7.5], 1 mM EDTA). Incubations were carried out at 48°C for the 5′ probe and 52°C for the 3′ probe for 18 h and were terminated by the addition of 1,000 U of S1 nuclease in a standard buffer. The S1 nuclease digestion was carried out at 37°C for 1 h. The samples were then ethanol precipitated and electrophoresed on an 8 M urea–5% acrylamide gel. The gels were dried on Whatman 3MM paper and exposed to Kodak XAR-5 X-ray film with an image intensifier overnight at −70°C.

RESULTS

Characterization of the mouse ζ-globin gene. A physical linkage map of the functional mouse α-globin genes was previously derived with cloned DNA fragments derived from a BALB/c mouse (see map; Fig. 1) (23). The 5′-most gene (established by data presented below) is the embryonic gene, which we refer to as ζ to comply with nomenclature established for the analogous human gene. Although the embryonic gene, which is located ca. 6 kb 5′ of the adult α_1-globin gene, is clearly an α-like gene (total coding nucleotide homology, 62%) (see below), it did not cross-hybridize to probes derived from the adult α_1-globin gene even when genomic DNA blots were washed at low stringency (40°C). Ultimately, the mouse ζ gene was isolated with a human ζ cDNA clone as the probe (kindly provided by B. Forget). In retrospect, having determined the structure of the ζ-globin gene (see below), we can see that strong homology between embryonic and adult genes is limited to a small sequence that corresponds to the end of the second and the beginning of the third exons. This region, which retains 85% homology between embryonic and adult sequences, is important in its interaction with the β-globin chain as well as its binding to iron.

The complete 1,962-base sequence of the ζ-globin gene, including portions of the 5′ and 3′ flanking regions and the two intervening sequences, is shown in Fig. 2. From this sequence we can deduce the 142-amino acid sequence of the mouse embryonic ζ-globin chain, including the N-terminal methionine residue which is cleaved from the mature form of the polypeptide (38). The deduced amino acid sequence is indicated beneath the nucleotide sequence shown in Fig. 2. A transcriptional initiation site can be inferred at position 430, ca. 30 base pairs (bp) 3′ to a TATAAA sequence. The canonical [CCAAT] box (4, 15) is altered to [CCAGT] at position 364 (i.e., 66 bases upstream from the initiation site). The same sequence is found in the goat ζ gene (S. Wernke and J. Lingrel, unpublished data). The three exons are bordered by an appropriate splice donor or acceptor site. A polyadenylate addition signal (AATAAA) (32) is located between positions 1920 and 1925.

Two features of the gene structure are of particular interest by virtue of their comparison with the human gene (33). In both cases, the first intervening sequence is a long one by comparison with adult α- and β-globin genes and contains an unusual sequence. The human ζ-globin gene contains the sequence ACAGTGGGGAGGGG repeated 12 times. These sequences are absent in the mouse. On the other hand, the dinucleotide GT is tandemly repeated 22 times (positions 1007 to 1051) in the mouse ζ-globin gene and is immediately followed by a GA repeated 9 times (positions 1054 to 1071). Neither sequence occurs as a repeat within the human gene.

To detect regions of conserved homology that lie outside the immediate coding regions of the mouse and human ζ-globin genes, a heteroduplex structure was formed between subcloned DNA fragments that include both coding and flanking segments of the mouse and human genes. An example of such a molecule is shown in Fig. 3. It reveals expected regions of homology that include the coding segments of the genes and a region of divergence that includes the large, first intervening sequence. Surprisingly, however,

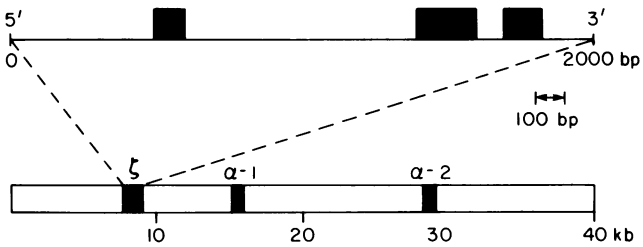


FIG. 1. Physical map of the α-globin locus of the BALB/c mouse. The three α-like globin genes are represented by solid regions and identified in the figure. The map was established by the comparison of overlapping segments of DNA cloned in phage λ from genomic DNA of the BALB/c mouse (23). The expanded representation of the embryonic gene (ζ) is indicated above the larger map, wherein the solid boxes represent the three ζ exon sequences. This structure and orientation is established by the sequence data presented in Fig. 2.

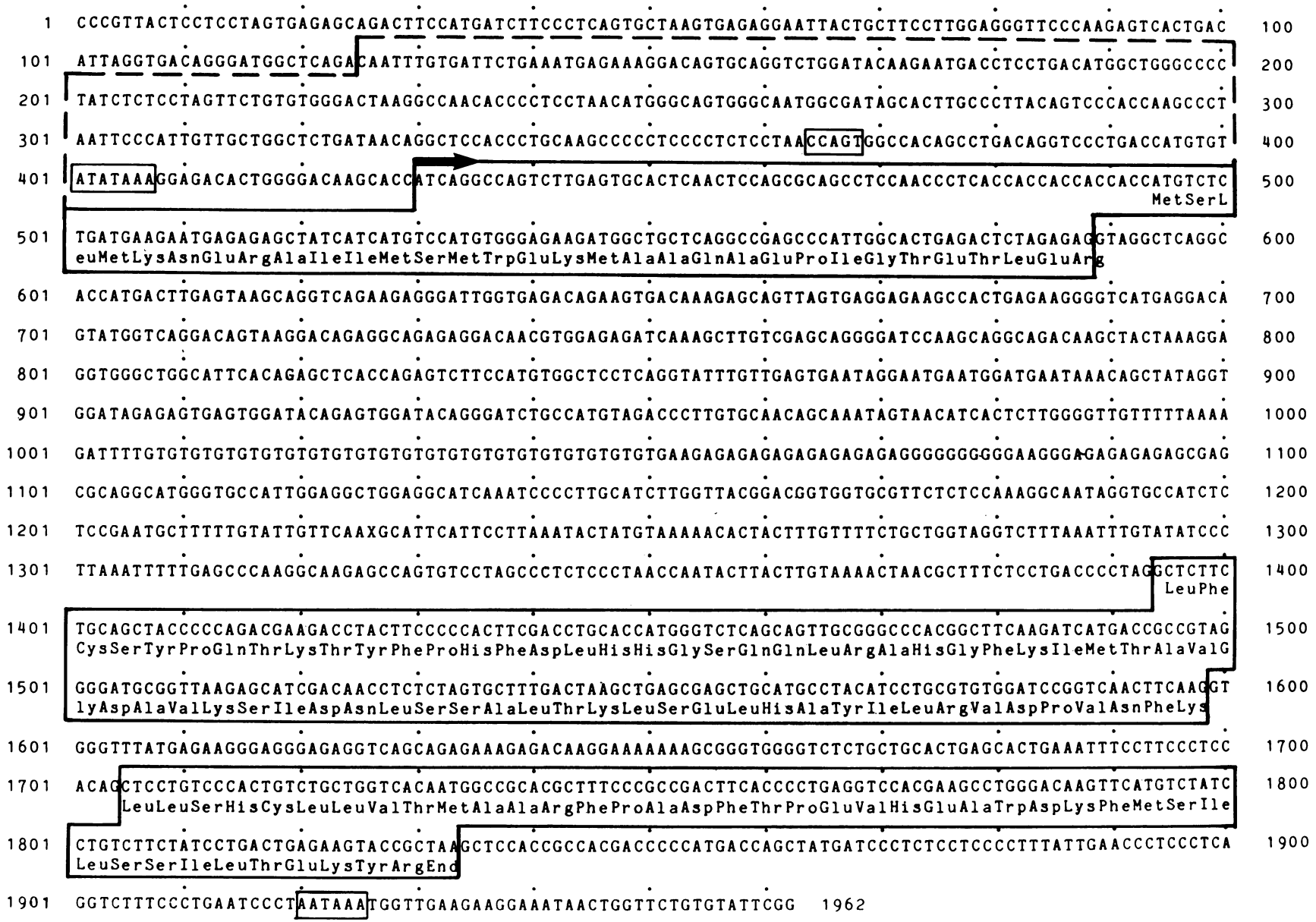


FIG. 2. The complete sequence of the embryonic ζ-globin gene of the BALB/c mouse. The three portions of the sequence enclosed within the solid boxes represent the three ζ exons. The deduced amino acid sequence of the ζ-globin protein is shown beneath the gene sequence. The arrow indicates the predicted transcription initiation site at position 430 with two small boxed sequences, one containing a TATA sequence and another containing an analog of the canonical CCAAT sequence here altered to CCAGT, 5′ to the initiation site. The conserved polyadenylate addition signal at position 1920 is also enclosed within a solid box. The dashed box enclosing ca. 300 bases at the 5′ portion of the gene is a sequence that is highly conserved between the mouse and the human embryonic ζ genes.

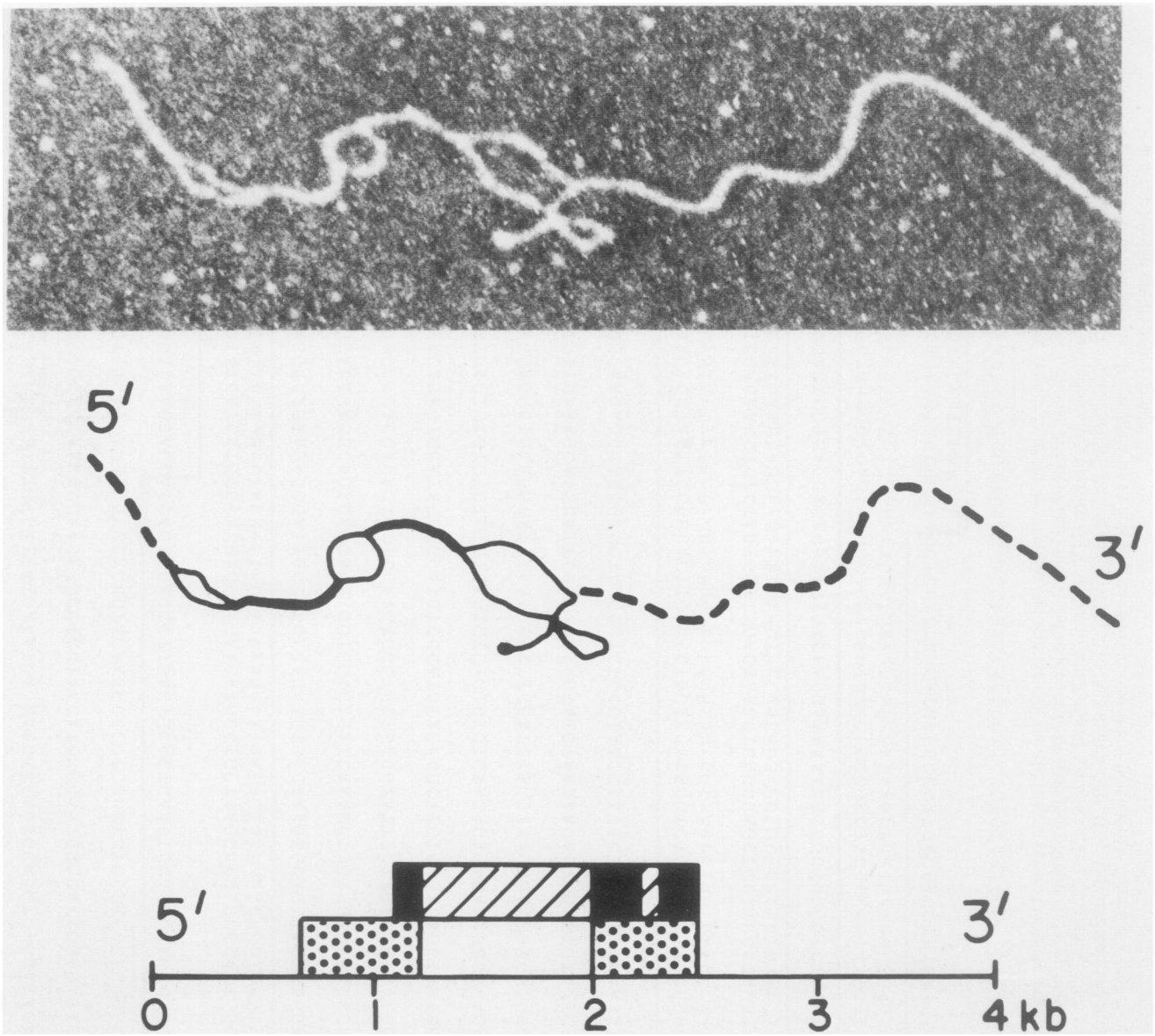


FIG. 3. A comparison of the mouse and human embryonic ζ genes by heteroduplex analysis. The electron micrograph shows a heteroduplex molecule formed by annealing fragments of the mouse (*Eco*RI) and human (*Hin*dIII) ζ genes subcloned in the same orientation into the *Eco*RI and *Hin*dIII sites of the plasmid pBR322, respectively. The plasmids have been linearized by digestion at a unique *Sal*I site in pBR322. The tracing represents the molecule shown above in which solid, thick lines represent mouse and human annealed sequences; solid, thin lines represent unannealed mouse and human sequences; and dashed lines represent pBR322 sequences. The diagrammatic map at the bottom of the figure indicates the extent of heteroduplex formation and mouse and human homology (dotted boxes) compared with the structure of the mouse ζ-globin gene (above), in which exons are represented as solid boxes and intervening sequences as hatched boxes.

it also reveals a long (ca. 300-bp) region of conserved sequence that extends beyond the transcriptional initiator site and promoter sequence within the 5′ flanking region of both ζ-globin genes (Fig. 3). By contrast, no such extensive 5′ homology is found between the adult α-globin genes of mouse and human. The extent and character of this conserved region is shown by a direct comparison of the two sequences in this region (Fig. 4). The region retains ca. 70% homology between the two sequences as compared with an 80% homology between the two amino acid coding regions. It is also worth noting that the stem and loop structure noted in the heteroduplex molecule (Fig. 3) is present in the human gene but absent in that of the mouse.

Expression of the ζ-globin gene in yolk sac cells. The yolk sac makes large amounts of the ζ-globin polypeptide. As far as we know, this is the only organ that produces it, and it does so only for a few days. From our cloned mouse ζ-globin gene we made specific 5′ and 3′ probes (see Fig. 5 and 6) with which we could study the transcription of the gene in the yolk sac. Using total RNA from the yolk sac of 10- and 12-day-old mouse embryos, we determined the sites of transcriptional initiation and polyadenylation of ζ mRNA by S1 nuclease analysis. A 1.7-kb *Bam*HI fragment containing the first exon and ca. 800 bp of 5′ flanking DNA was cloned into M13 which was used as a template to make radioactively labeled, single-stranded probes by primer extension.

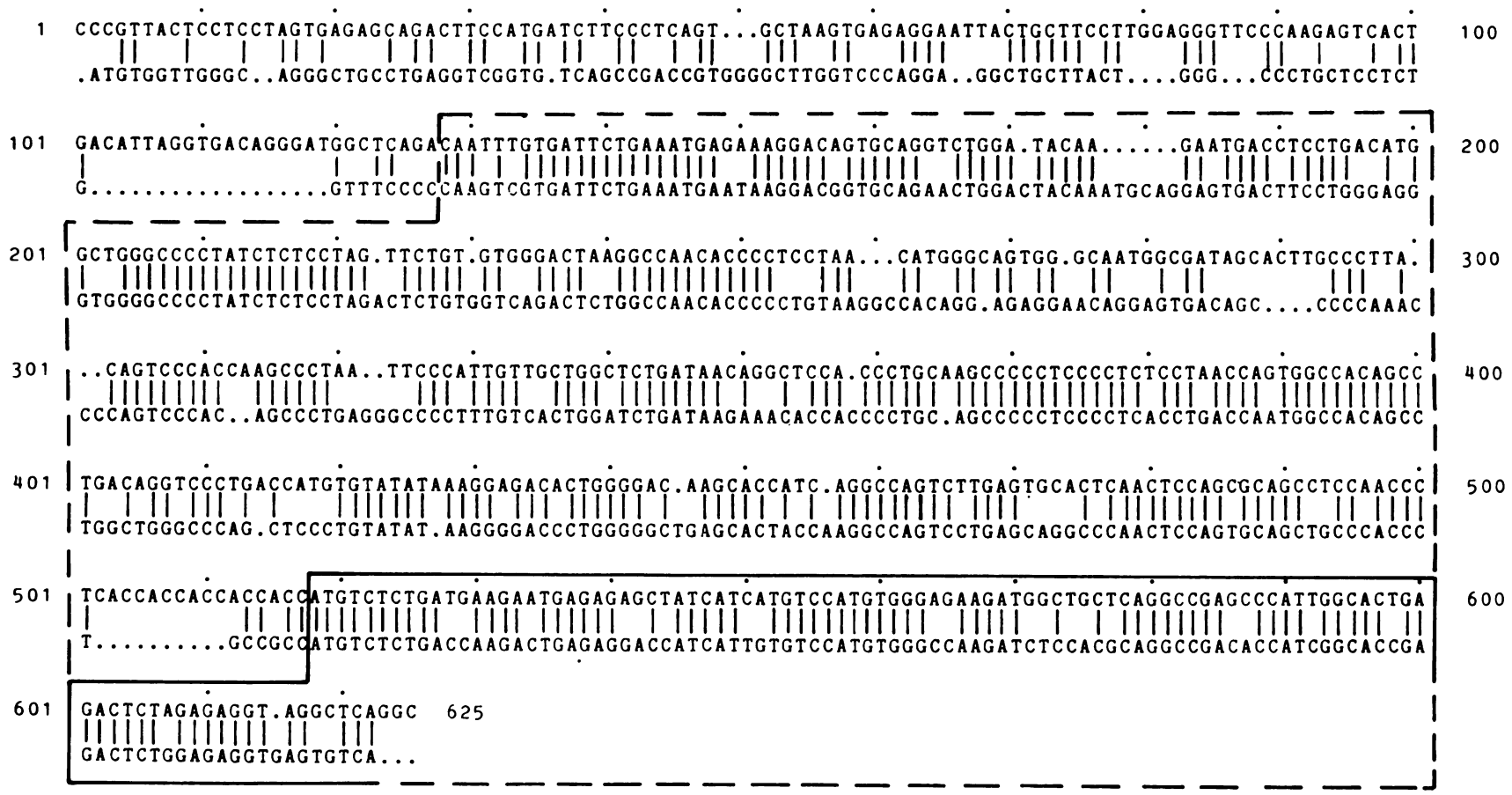


FIG. 4. Comparison of the sequences of the 5′ portions of the mouse and human ζ-globin genes. The top sequence is that of the mouse gene, and the bottom is that of the human gene. The coding sequence begins at nucleotide 518. The region 5′ homology is enclosed within a dashed line. Identical bases are indicated by vertical lines.

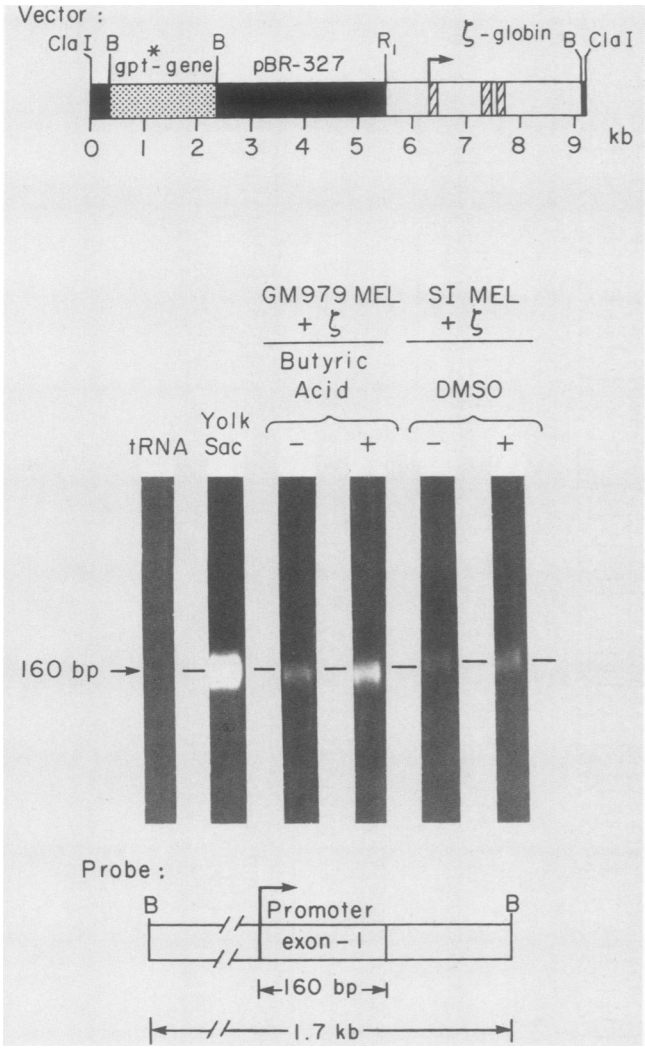


FIG. 5. Expression of the 5′ portion of the ζ-globin gene in yolk sac and permanently transfected mouse erythroleukemia (MEL) cells and response to agents that cause erythrodifferentiation. The upper diagram is a map of the transfecting vector containing the selectable *gpt* gene, pBR327 sequences, and cloned ζ-globin gene. The transcriptional initiation site of the ζ-globin gene is indicated by an arrow, and its exon sequences are indicated by hatched boxes. The middle portion of the figure shows the results of S1 nuclease protection analyses of mRNAs derived from the indicated cells with or without butyric acid (1.2 mM) or DMSO (1.5%) treatment. Untransfected control mouse erythroleukemia cells (see Fig. 6) failed to show any detectable ζ-globin fragment protection. The bottom portion of the figure shows a map of the ζ-globin gene DNA fragment used for the S1 nuclease protection assay. The expected position of the accurately transcribed, protected fragment (160 bases) is shown. B, Restriction site for *Bam*HI; R_1, restriction site for *Eco*RI.

As can be seen in Fig. 5 (lane 2), a very prominent band of 160 nucleotides of the probe was protected by the mouse yolk sac RNA. This maps the site of transcriptional initiation of the ζ-globin gene to position 430, ca. 30 bp downstream to a TATAAA sequence (Figure 2). Similarly, a 1.2-kb *Bam*HI fragment containing exon 3 and at least 800 bp of DNA downstream of the putative polyadenylation site was cloned into M13 and used to make a specific 3′ S1 probe. Figure 6 (lane 2) shows clearly that a band of 240 nucleotides was protected by the yolk sac RNA, mapping the polyadenylation site ca. 20 nucleotides 3′ to the AATAAA signal.

Expression of ζ-globin gene in transfected cell lines. One of the major limitations in studying the regulated expression of the mouse ζ-globin gene has been the absence of an established cell line in which this gene can be reliably expressed. Such a line would overcome the difficulty of obtaining significant numbers of nucleated, embryonic erythrocytes, the only cell in which this gene is normally expressed. To overcome this problem and to establish a system in which some of the requirements for ζ-globin gene expression could be established, we turned to mouse erythroleukemic cell lines. These lines can be induced to undergo a form of erythrodifferentiation in which both the adult α- and β-globin genes are strongly induced (35). In the two cell lines we have used, GM979 and SI, both DMSO and butyric acid (22) produce this effect, but neither reliably induces expression of the endogenous ζ-globin genes in either of the cell lines. (Occasionally, the GM979 cell line will express very low levels of ζ-globin mRNA after treatment with butyric acid [see Fig. 6].)

In the absence of a cell line that reliably expresses the endogenous ζ-globin gene, either with or without induction, we constructed a hybrid plasmid that carries the mouse ζ-globin gene in association with a selectable marker for the *E. coli* enzyme, guanosine phosphoribosyltransferase (*gpt*) (29) (see Fig. 5). Thus, we could permanently introduce the gene into these cell lines and determine whether the embryonic gene could be expressed in an adult erythropoietic cell background. For comparison, we selected an immunoglobulin-producing B-cell line (M12) as an example of a non-erythroid hematopoietic cell (20).

Using selection in conjunction with transfection by electroporation (31), we were able to isolate several transfectant clones from each cell line. This technique generally introduces few gene copies per cell (31, 37). In our case (data not shown), the copy number was less than five. The ability of these clones to express the ζ-globin gene was measured by the S1 nuclease protection assay, making use of probes derived from both the 5′ and 3′ portions of the mouse ζ gene (see Fig. 5 and 6, respectively). Although only one transfectant analysis is shown for each cell line, several transfectants were assayed in each case with comparable results. The exogenous ζ-globin gene was expressed in each cell line, including the B-cell line (see Fig. 6), and its transcripts protected DNA fragments identical to those protected by yolk sac mRNA, indicating accurate initiation (Fig. 5) and polyadenylation (Fig. 6) at the predicted sites. Moreover, treatment with DMSO and butyric acid appeared to increase the amount of ζ-globin gene transcript in all three cell lines, the effect being most pronounced in the case of the SI transfectant cell line induced with DMSO (Fig. 6). These results have been supported by direct mRNA gel analyses which, in most cases, show the transfectant cells to be producing normal-sized ζ-globin mRNA (data not shown). In a few cases, larger transcripts were seen that may be the result of aberrant splicing. In separate experiments (data not shown), we have determined the level of *gpt* gene transcript and find no appreciable difference with or without DMSO or butyric acid treatment.

Structures of the mouse and human ζ-globin genes in a test of mouse and human evolutionary lineage. In addition to providing a structural definition of the mouse ζ-globin gene and protein, our sequence data provide a link that allows us to estimate the evolutionary relationship that exists between several higher vertebrate species. The embryonic and adult

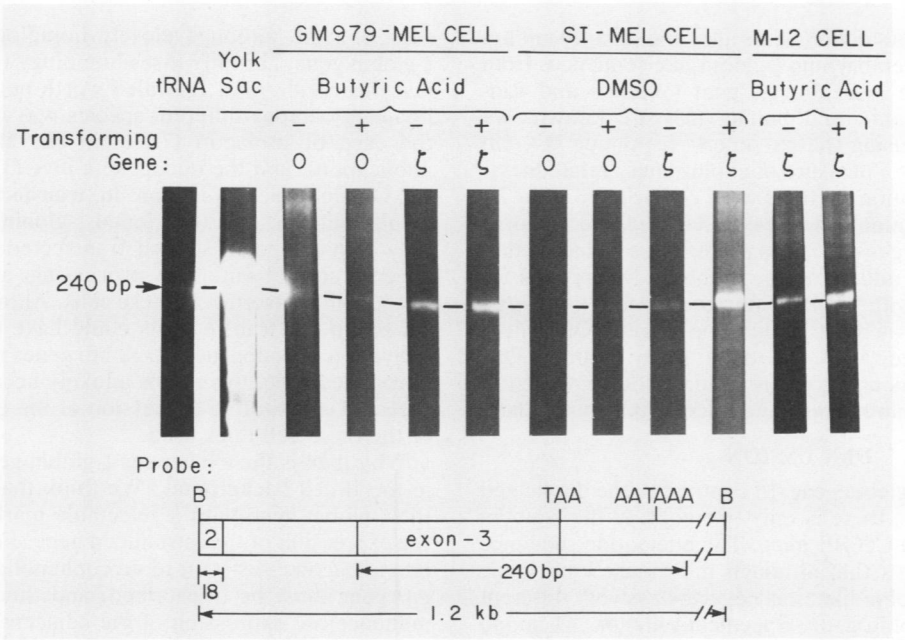
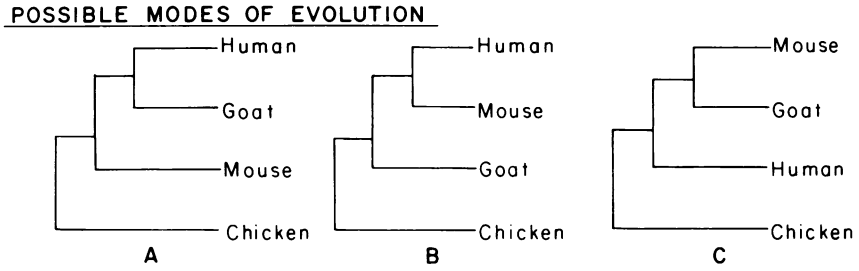


FIG. 6. Expression of a permanently transfected ζ-globin gene in cultured mouse erythroleukemic (MEL) and B-cells and its response to agents that induce erythrodifferentiation. The two mouse erythroleukemic cell lines are identified above their respective lanes. The M-12 line is a mouse B-cell line. The indicated cells (with the exception of the yolk sac controls) permanently transfected with the *gpt*-ζ-globin gene vetor indicated in Fig. 5. The S1 nuclease protection assay was carried out with a fragment of DNA derived from the 3′ terminus of the ζ-globin gene that is indicated at the bottom of the figure. A fragment of 240 bases should be protected by a normally polyadenylated ζ-globin gene transcript. The plus and minus symbols above certain lanes indicate whether or not the cells have been treated by the indicated agent. The "O" indicates that the cells analyzed do not contain a transfecting plasmid. B, Restriction site for *Bam*HI.

α-globin genes evolved from a common ancestral gene ca. 400 million years ago, before mammals and birds diverged (11, 33). These numbers are based on the fact that there is a strong correlation between the relatedness of species and the similarity in amino acid and DNA sequence. For example, changes between adult α-globin mouse sequences and adult α-globin chicken sequences are fewer than the changes between adult α mouse sequences and mouse ζ sequences. This suggests that the duplication that created the adult and embryonic α lineages occurred before mammals and birds diverged.

Comparison of amino acid sequences (12) provides an approximate guide to divergence dating among species and suggests that birds began to diverge from humans 300 million



PARSIMONY ANALYSIS

	A	B	C
First exon	54	60	59
Second exon	100	104	104
Third exon	71	84	83
TOTALS	225	248	246

FIG. 7. Potential modes of evolution as predicted by parsimony analysis of the amino acid sequences of the human, mouse, goat, and chicken ζ-like globin genes. The upper portion of the figure shows three possible modes of evolution between the four species in which a close evolutionary relationship is indicated by the shortest connection. The bottom portion of the figure (parsimony analysis) shows the minimum number of base changes required to relate the exons of each species to a common ancestor according to the three modes of evolution. Mode A requires the fewest changes, i.e., is the most parsimonious.

years ago, goats ca. 80 million years ago, and mice 70 million years ago. Since the embryonic ζ-globin-like sequences from chicken (16), mouse (see above), goat (Wernke and Lingrel, unpublished data), and human (33) are known, we compared them by using the parsimony technique (17, 18) which can be used as a measure of evolutionary relatedness. Three possible evolutionary trees were constructed (Fig. 7A to C). For each amino acid, we calculated the minimal number of base changes needed to produce the variation that is seen today. After adding up the number of mutations for each tree, the tree with the least number of mutations—the most parsimonious—is most likely to be correct. Our calculations point to tree A as the most likely option. This suggests that humans and goats are more closely related in evolutionary time than humans and mice or mice and goats.

DISCUSSION

Structure of the ζ-globin gene. In contrast to the duplicated adult α-globin genes, there is only one copy of the ζ-globin gene in BALB/c and C56BL mice. The nucleotide sequence presented here shows that although the mouse embryonic ζ-globin gene is α-globin-like, it is nevertheless very different from the adult α-globin gene. Specifically, 66 of 141 amino acids differ. In addition, the first intron, which is characteristically short (100 to 200 bases) in all adult α- and β-globin genes, is over 800 bases long in the ζ-globin gene.

Comparison of these genes also allows us to comment on the functional value of some of their more remarkable features. For example, Proudfoot et al. (33) have reported the existence of repetitive sequences within the first intervening sequence of the human ζ-globin gene, speculating on their functional significance. However, these sequences are not found in the mouse ζ-globin gene, although the mouse has a different set of repetitive sequences within its first intervening sequence. The fact that neither of these sequences are conserved between the species suggests, but does not prove, that they may be of little functional value. The GT repeat found in the first intron of the mouse ζ-globin gene (Fig. 2) is a sequence that can partially exist in the Z-DNA configuration (19, 34), and therefore its presence within an active gene makes it an attractive subject for functional speculation (19). On the other hand, its absence in the human gene makes this property difficult to evaluate.

Comparison of this sequence with the corresponding ζ genes from human, goat, and chicken shows clearly that there is more homology between the embryonic genes of different species than there is between embryonic and adult α-globin genes of the same species. This confirms the notion that the duplication of the ancestral gene that produced the adult and embryonic genes occurred long before human, goat, mouse, and chicken ancestral species diverged. The sequence also enables us to add another point to the evolutionary tree which relates species via their globin genes. Our parsimony analysis points to an evolutionary relationship that places goats and humans closer to one another than to mice. This analysis agrees with a phylogenetic tree devised by Czelusniak et al. (11) which was based on the analysis of 40 globin genes.

The differential regulation of the ζ-globin gene is likely to depend in part upon *cis*-acting elements encoded near or within the locus. In this respect, it is particularly noteworthy that a 300-bp sequence is conserved between the mouse and human genes in the 5′ flanking region. Such homologous segments, especially when they are conserved between species that have been apart for 70 million years, are likely to be of functional significance. Interestingly, the pseudo-ζ gene of man, although closely homologous to the normal ζ-globin gene, abruptly loses homology to the active gene in just this region (33). It is also worth noting that 5′ flanking homology among different species was recently reported in the case of prolactin (7), actin (D. Jaffe, submitted for publication), and the oncogene c-*myc* (3).

ζ-Globin gene expression in transfected cell lines. Our results indicate that the cloned ζ-globin gene can be faithfully transcribed in several transfected cell lines; i.e., we observe correct initiation, processing, and polyadenylation of ζ-globin transcripts in these cells. Although the transcripts we see in our transfectants could have originated from the activation of endogenous ζ-globin genes rather than from the introduced gene, this seems unlikely because we have never seen the constitutive expression of the ζ-globin gene in any of the three cell lines used.

Why then is the exogenous ζ-globin gene expressed, even in the B-cell background? We think that the association of the ζ-globin gene with a selectable marker helps to ensure the expression of the introduced gene. For a transfected cell line to become resistant to mycophenolic acid, the incoming *gpt* gene must be transcribed, and this requirement could influence the expression of the adjacent ζ-globin gene. The transfecting gene may have integrated into a transcriptionally active site or may be active because it happens to be in a transcriptionally competent configuration (e.g., unmethylated) compared with the bulk of the chromatin. It is also possible that the ζ-globin gene is influenced by the simian virus 40 enhancer sequence associated with the promoter driving the *gpt* gene in the transfecting plasmid. Nevertheless, the expression of the ζ-globin gene in transfected B-cells is surprising in view of the presumed requirement for a B-cell-specific enhancer for the expression of adult globin genes in this cell background (1). Although the ζ-globin gene expression that we observe in the B-cell may reflect the fortuitous integration of the transfecting plasmid in the vicinity of a B-cell-active enhancer region or some special property of the embryonic globin gene, this observation suggests that tissue-specific restriction of expression may be less stringent than originally supposed, at least in the transfected cell system.

The fact that the transfecting ζ-globin gene is inducible in hematopoietic cells is encouraging in that it provides a system in which *cis*-acting elements required in the expression and induction of the ζ-globin gene can be evaluated. In this regard, one of the most interesting features of the mouse ζ-gene sequence is the 5′ flanking region that is homologous to the 5′ flanking region of the human ζ gene. Gene transfer experiments focusing on this region should allow us to evaluate its role in regulating the expression of the ζ-globin gene and thereby establish a basis for its evolutionary conservation.

ACKNOWLEDGMENTS

We thank Hunt Potter for performing the electron microscopic analyses used in this study and for guiding us in the use of the electroporation technique, James Battey for helping us with computer analyses, and Allan Wilson for his helpful discussion regarding the parsimony analysis. We also thank Patrick Charny for providing the human ζ gene used in the heteroduplex study, Suzanne Wernke and Jerry Lingrel for graciously making available to us their unpublished sequence data regarding the embryonic goat gene, and Terri Broderick for expert assistance in the preparation of this manuscript.

Portions of this work were supported by grants from the American Business for Cancer Research Foundation and E. I. du Pont de

Nemours & Co., Inc. L.W. was supported by a grant from the North Atlantic Treaty Organization.

LITERATURE CITED

1. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
2. **Barker, J. E.** 1968. Development of mouse hematopoetic system. Dev. Biol. **18**:14–29.
3. **Battey, J., C. Moulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, and P. Leder.** 1983. The human c-myc oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell **34**:779–787.
4. **Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon.** 1980. The ovalbumin gene—sequence of putative control regions. Nucleic Acids Res. **8**:127–142.
5. **Berk, A., and P. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
6. **Brotherton, T. W., D. H. K. Chui, J. Gauldie, and M. Patterson.** 1979. Hemoglobin ontogeny during normal mouse fetal development. Proc. Natl. Acad. Sci. U.S.A. **76**:2853–2857.
7. **Camper, S. A., D. N. Luck, Y. Yao, R. P. Woychik, R. G. Goodwin, R. H. Lyons, Jr., and F. M. Rottman.** 1984. Characterization of the bovine prolactin gene. DNA **3**:237–249.
8. **Chirgwin, J., A. Aeybyle, R. McDonald, and W. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
9. **Craig, M. L., and E. S. Russel.** 1964. A developmental change in hemoglobins correlated with an embryonic red cell population in the mouse. Dev. Biol. **10**:191–201.
10. **Cudennec, C. A., J.-P. Thiery, and N. M. Le Douarin.** 1981. In vitro induction of adult erythropoiesis in early mouse yolk sac. Proc. Natl. Acad. Sci. U.S.A. **78**:2412–2416.
11. **Czelusniak, 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 hemoglobin genes. Nature (London) **298**:297–300.
12. **Dayhoff, M. O.** 1972. Atlas of protein sequence and structure, vol. 5. National Biomedical Research Foundation, Washington, D.C.
13. **Dodgson, J. B., K. C. McCune, D. J. Rusling, A. Krust, and J. D. Engel.** 1981. Adult chicken α-globin genes, α^A and α^D: no anemic shock α-globin exists in domestic chickens. Proc. Natl. Acad. Sci. U.S.A. **78**:5998–6002.
14. **Dodgson, J. B., and J. D. Engel.** 1983. The nucleotide sequence of the adult chicken α-globin genes. J. Biol. Chem. **258**: 4623–4629.
15. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21**:635–668.
16. **Engel, J. D., D. J. Rusling, K. C. McCune, and J. B. Dodgson.** 1983. Unusual structure of the chicken embryonic α-globin gene, Π'. Proc. Natl. Acad. Sci. U.S.A. **80**:1392–1396.
17. **Farris, J. S.** 1967. Definitions of taxa. Syst. Zool. **16**:174–175.
18. **Ferris, S. D., A. C. Wilson, and W. M. Brown.** 1981. Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. Proc. Natl. Acad. Sci. U.S.A. **78**: 2432–2436.
19. **Gilmour, R. S., D. A. Spandidos, J. K. Vass, J. W. Gow, and J. Paul.** 1984. A negative regulatory sequence near the mouse β-major globin gene associated with a region of potential Z-DNA. EMBO J. **3**:1263–1272.
20. **Glimcher, L. H., K.-J. Kim, I. Green, and W. E. Paul.** 1982. Ia antigen-bearing B cell tumor lines can present protein antigen and alloantigen in a major histocompatibility complex-restricted fashion to antigen-reactive T cells. J. Exp. Med. **155**:445–459.
21. **Lauer, J., C.-K. Shen, and T. Maniatis.** 1980. The chromosomal arrangement of human α-like globin genes: sequence homology and α-globin gene deletions. Cell **20**:119–130.
22. **Leder, A., and P. Leder.** 1975. A potent inducer of erythroid differentiation in cultured erytholeukemic cells. Cell **5**:319–322.
23. **Leder, A., D. Swan, F. Ruddle, P. D'Eustachio, and P. Leder.** 1981. Dispersion of α-like globin genes of the mouse to three different chromosomes. Nature (London) **293**:196–200.
24. **Leder, P., D. Tiemeier, and L. Enquist.** 1977. EK2 derivatives of bacteriophage λ useful in the cloning of DNA from higher organisms: the λ gtWES system. Science **196**:175–177.
25. **Ley, T. J., N. P. Anagnou, G. Pepe, and A. W. Nienhuis.** 1982. RNA processing errors in patients with β-thalassemia. Proc. Natl. Acad. Sci. U.S.A. **79**:4775–4779.
26. **Liebhaber, S. A., M. Goosens, and Y. W. Kan.** 1981. Homology and concerted evolution at the α1 and α2 loci of human α-globin. Nature (London) **290**:26–29.
27. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
28. **Michelson, A. M., and S. H. Orkin.** 1980. The 3′ untranslated regions of the duplicated human α-like globin genes are unexpectedly divergent. Cell **22**:371–377.
29. **Mulligan, R. C., and P. Berg.** 1981. Selection for animal cells that express the Escherichia coli gene coding for xanthine-guanine phosphoribosyltransferase. Proc. Natl. Acad. Sci. U.S.A. **78**:2072–2076.
30. **Nishioka, Y., and P. Leder.** 1979. The complete sequence of a chromosomal mouse α-globin gene reveals elements conserved throughout vertebrate evolution. Cell **18**:875–882.
31. **Potter, H., L. Weir, and P. Leder.** 1984. Enhancer-dependent expression of human κ immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. Proc. Natl. Acad. Sci. U.S.A. **81**:7161–7165.
32. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′ noncoding region sequences in eukaryotic messenger RNA. Nature (London) **263**:211–214.
33. **Proudfoot, N. J., A. Gil, and T. Maniatis.** 1982. The structure of the human ζ-globin gene and a closely linked, nearly identical pseudogene. Cell **31**:553–563.
34. **Rich, A.** 1982. Right-handed and left-handed DNA: conformational information in genetic material. Cold Spring Harbor Symp. Quant. Biol. **47**:1–12.
35. **Ross, J., H. Aviv, and P. Leder.** 1973. Characterization of synthetic rabbit globin DNA. Arch. Biochem. Biophys. **158**:494–502.
36. **Tiemeier, D., S. Tilghman, and P. Leder.** 1977. Purification and cloning of a mouse ribosomal gene fragment in phage λ. Gene **2**: 173–191.
37. **Wong, T.-K., and E. Neumann.** 1982. Electric field mediated gene transfer. Biochem. Biophys. Res. Comm. **107**:584–587.
38. **Yoshida, A., and W. Lin.** 1972. NH_2-terminal formylmethionine- and NH_2-terminal methionine-cleavage in rabbits. J. Biol. Chem. **247**:952–957.