A Molecular and Evolutionary Study of the β-Globin Gene Family of the Australian Marsupial *Sminthopsis crassicaudata*

S.J.B. Cooper, *† R. Murphy,* G. Dolman,* D. Hussey, *and R.M. Hope*

*Department of Genetics, University of Adelaide, Australia; and †Evolutionary Biology Unit, South Australian Museum*

P-globin gene families in eutherians (placental mammals) consist of a set of four or more developmentally regulated genes which are closely linked and, in general, arranged in the order 5′-embryonic/fetal genes-adult genes-3′. This cluster of genes is proposed to have arisen by tandem duplication of ancestral P-globin genes, with the first duplication occurring 200–155 MYBP just prior to a period in mammalian evolution when eutherians and marsupials diverged from a common ancestor. In this paper we trace the evolutionary history of the P-globin gene family back to the origins of these mammals by molecular characterization of the β-globin gene family of the Australian marsupial *Sminthopsis crassicaudata*. Using Southern and restriction analysis of total genomic DNA and bacteriophage clones of P-like globin genes, we provide evidence that just two functional P-like globin genes exist in this marsupial, including one embryonic-expressed gene (Xc-e) and one adult-expressed gene (S_c-β), linked in the order 5′-e-β-3′. The entire DNA sequence of the adult β-globin gene is reported and shown to be orthologous to the adult P-globin genes of the North American marsupial *Didelphis virginiana* and eutherian mammals. These results, together with results from a phylogenetic analysis of mammalian P-like globin genes, confirm the hypothesis that a two-gene cluster, containing an embryonic- and an adult-expressed P-like globin gene, existed in the most recent common ancestor of marsupials and eutherians. Northern analysis of total RNA isolated from embryos and neonatals indicates that a switch from embryonic to adult gene expression occurs at the time of birth, coinciding with the transfer of the marsupial from a uterus to a pouch environment.

**Introduction**

In mammals, the major oxygen-transporting molecule hemoglobin is composed of two α- and two β-globin polypeptides, with each encoded by families of α-like and P-like globin genes which are differentially expressed during development. Molecular genetic studies have revealed that the individual member genes of each family have arisen through a series of tandem duplications of ancestral α-like and P-like globin genes during the evolution of higher vertebrates (Jeffreys et al. 1980; Czelusniak et al. 1982; Goodman et al. 1984; Hardison 1984). In particular, molecular studies of β-globin gene families in species from four orders of eutherian mammals (Lagomorpha, Artiodactyla, Primates, and Rodentia) have provided evidence that a four- to five-gene ancestral β-globin cluster existed prior to the radiation of these eutherian mammals (Goodman et al. 1984; Hardison 1984). The first duplication of a single ancestral P-globin gene is estimated to have occurred between 200 and 155 MYBP and resulted in the production of two genes, one of which became developmentally delayed in its expression while the other became expressed in embryonic tissues only (Czelusniak et al. 1982; Efstratiadis et al. 1980). Prior to the eutherian radiation, between 100 and 50 MYBP, the embryonic gene is proposed to have duplicated a further two times and the adult gene duplicated once, producing a P-globin gene cluster consisting of three embryonic genes (ε, γ, and η) and two adult genes (6 and β) linked in the order:

5′-ε-γ-η-δ-β-3′ (Goodman et al. 1984). The timing of these later duplications coincides with the divergence of marsupials and eutherians from a common ancestor, approximately 155–100 MYBP (Air et al. 1971; Hope, Cooper, and Wainwright 1990). Molecular characterization of β-globin gene families in marsupials, therefore, provides a critical test of these hypotheses. Comparisons between marsupial and eutherian β-globin gene families allow the history of the gene family to be traced back to the origin of these major mammalian infrclasses.

Molecular analysis of the β-globin gene family of the American marsupial *Didelphis virginiana* (Koop and Goodman 1988) and the Australian dasyurid marsupial *Sminthopsis crassicaudata* (Cooper and Hope 1993) confirmed that an embryonic P-like globin gene existed in marsupials and was already differentiated, with respect to its developmental expression, prior to divergence of the eutherian and marsupial lineages. In addition, Koop and Goodman (1988) provided evidence that just two genes existed in *D. virginiana*, including one adult-expressed gene and one gene which was orthologous to eutherian embryonic P-like globin genes. Although they were unable to confirm the physical linkage and chromosomal arrangement of the two genes, their results suggested that an ancestral P-globin cluster consisting of only two genes was present in the most recent common ancestor of marsupials and eutherians. *D. virginiana* and *S. crassicaudata* belong to separate cohorts (Ameridelphia and Australidelphia, respectively) within the infraclass Marsupialia (Aplin and Archer 1987). Each of these cohorts is estimated to have diverged from a common ancestor approximately 105 MYBP (Richardson 1988; Hope, Cooper, and Wainwright 1990). During a similar time span in eutherian mammals there were considerable alterations in hemoglobin gene number and expression through processes such as gene duplication (Townes, Fitzgerald, and Lin-
gene conversion (Jeffreys et al. 1982; Scott et al. 1984), gene inactivation (Hardies, Edgell, and Hutchison 1984) and recruitment of genes into different developmental regimes (Goodman et al. 1987; Tagle et al. 1988). It is therefore possible that more than two progenitors of the five eutherian B-like globin loci existed prior to the divergence of the marsupial and eutherian lineages but by gene conversion and inactivation only two genes remained in the lineage leading to D. virginiana.

Different molecular forms of hemoglobin, restricted in the timing of their expression to neonatal (pouch young) and embryonic tissues, have been detected from a number of Australian macropodid marsupials (Richardson and Russell 1969; Holland et al. 1988), a possum, Trichosurus vulpecula (Hope 1970; Calvert, Holland, and Hinds 1993), and a dunnart, Sminthopsis crassicaudata (Holland et al. 1994). The presence of these hemoglobins suggests that Australian marsupials may have a more complex set of B-globin genes than D. virginiana. Although embryogenesis occurs in the uterus of both marsupials and eutherians, marsupials are born at a comparatively early stage of embryological development and their fetal growth phase, unlike the longer intrauterine growth phase of eutherians, occurs in a pouch environment that is hypoxic and hypercapnic compared to ambient air (Almeida and Roch 1932; Bailey and Dunnet 1960; Baudinette et al. 1988; Holland et al. 1988). It is therefore possible that some marsupials may have unique forms of hemoglobin specifically adapted to the oxygen transport requirements of pouch neonates. For all these reasons we have investigated the evolution and expression of B-like globin genes from the Australian marsupial Sminthopsis crassicaudata.

Materials and Methods

All animals were obtained from a pedigreed colony of Sminthopsis crassicaudata maintained in the Genetics Department, R. A. Fisher Laboratories, The University of Adelaide (Bennett et al. 1982, 1990). Probes were as follows: pDG-5R, a 184-bp Rso I fragment derived from a cDNA clone of adult β-globin mRNA (pDG-5) from the marsupial Dasyurus viverrinus, containing the terminal 3′ terminal 63 bp of exon 2 and all of exon 3, except the terminal 5 bp (Wainwright and Hope 1985); pDG-5HS, a 400-bp HindIII/Sac I fragment derived from pDG-5; and pSG-2H, a 467-bp Hind1I1 fragment derived from a subclone (pSG-2) of a genomic clone (λSG-3), which contains the terminal (3′) 12 bp from the first exon, the entire first intron (120 bp) and second exon (223 bp), and the first (5′) 112 bp of the second intron, of the embryonic B-like globin gene of S. crassicaudata (S.c-E). For further details of clones see Cooper and Hope (1993). An embryonic-specific probe for Northern analysis was obtained by PCR amplification of the first exon of S.c-E from the subclone pSG-2 using a primer containing the CCAAT-box in the 5′ flanking region (5′-CTCTGACCAATAGCTTCAG-3′) and a primer designed from first intron sequence (5′-GCTTCTTGGTCTTC-3′). PCR amplifications were carried out as specified below for sequence analysis with the exception that the reaction buffer contained 2.5 mM MgCl₂ and an annealing temperature of 56°C with 30 cycles of amplification was used.

Genomic DNA Isolation, Genomic Library Construction, and Screening

S. crassicaudata genomic DNA was prepared from 100 mg of liver tissue using a standard phenol/chloroform extraction procedure (Hope et al. 1992). Genomic DNA from one animal (1843.1 b) was digested partially with the restriction endonuclease Sau3A (Boehringer-Mannheim) and cloned into λ-GE M1 I (Promega) using procedures specified by Promega. Bacteriophage DNA from 6 X 10⁵ plaques was transferred to Hybond N+ (Amersham) filters using procedures specified by Amersham, and hybridized with an adult β-globin probe pDG-5R, radiolabeled with α²³P-dCTP to a specific activity of 3 X 10⁹ cpm/µg (Hodgson and Fisk 1987). Hybridization was carried out in a buffer containing 0.263 M Na₂HPO₄, 1 mM EDTA, 1% (w/v) BSA, 7% (w/v) sodium dodecyl sulphate (SDS) overnight at 65°C. Filters were washed in 1 X SSC (1 X SSC = 0.15 M NaCl, 15 mM trisodium citrate), 0.1% (w/v) SDS at 65°C, and autoradiographed for between 2 and 7 days at -70°C. Positively hybridizing clones were further purified by three rounds of low-density plating, plaque lifting, and screening. Phage DNA was purified from liquid lysates using the PEG (8000) procedure and small-scale isolation technique described in Sambrook, Fritsch, and Maniatis (1989) with minor modifications. After precipitation, phage DNA was solubilized, washed in 70% (v/v) ethanol, dried, and resuspended in TE (pH 8.0).

Southern Blotting and Probing

Genomic or phage DNA was digested with a range of restriction endonucleases, including BamHI, EcoRI, HindIII, and Pst I (Boehringer-Mannheim), in buffers supplied by the manufacturer. Digested DNA was electrophoresed on agarose gels in 1 X TAE buffer (Sambrook, Fritsch, and Maniatis 1989) and Southern transferred to Hybond N+ (Amersham) using procedures specified by Amersham. Probes were radio-labeled as described above and hybridization was carried out in a buffer containing 5 X SSPE (1 X SSPE = 0.15 M NaCl, 1 mM EDTA, 1mM Na₂HPO₄), 5 X Denhards, 0.1% (w/v) SDS, and 100 µg/ml single-stranded salmon DNA overnight at 65°C. Washing conditions were 65°C, 2 X SSC, 0.1% (w/v) SDS.

Sequence Analysis

Two strategies were adopted for DNA sequencing. One involved sequencing of plasmid(pGEM-3Zf(−)), (Promega) subclones containing globin-hybridizing regions using a PRISM® dye primer sequencing kit and primer extension from fluorescently labeled universal M 13 reverse and forward sequencing primers (ABI–Perkin Elmer). The second strategy involved a primer walking approach, to obtain sequence data from regions that were not subcloned and from both strands of the adult β-globin gene. The position and sequence of oligonucleotide primers is shown in Figure 1.
cleotide primers designed to span the entire adult β-globin gene are shown in figure 4. These primers were used to PCR-amplify regions of the gene in 50-μl reaction volumes containing 1 × reaction buffer (Promega), 4 mM MgCl₂, 0.2 mM each dNTP, 10 pm of each primer, and 50–100 ng of template DNA. PCR reactions were carried out on a Corbett (FTS320) thermocycler for 35 cycles (94°C, 45 s; 55°C, 45 s, and 72°C, 1–2 min) followed by a final incubation at 72°C for 6 min. PCR products were purified using glass-milk in a “BRESA-CLEAN” kit and sequenced using the PRISM® Ready Reaction Dyedexoxy terminator cycle sequencing kit (ABI–Perkin Elmer). Primers used for sequencing were the same as those used in the original PCR amplification. All reaction products were purified by phenol/chloroform extraction and ethanol precipitation (as specified by ABI) and sequenced on an ABI 373A DNA sequencer (Microbiology Department, University of Adelaide). Sequences were manually edited using the program SeqEdt (ABI).

Northern Analysis

RNA was isolated from an embryo (1 to 2 days before birth), and pouch young, between 0 and 36 days after birth, using Trizol “one step purification” reagent (Gibco BRL) and a procedure specified by BRL. RNA samples were denatured and electrophoresed on a formaldehyde agarose gel (1.4% w/v) agarose, 1 × MOPS buffer, 5% [v/v] formaldehyde) in 1 × MOPS buffer. RNA was Northern transferred to Hybond N+ (Amer- sham) as specified by Amersham. The Northern filter was hybridized with DNA probes using a procedure identical to that described above for Southern blots. Filters were washed in 1 × SSC, 0.1% (w/v) SDS at 65°C, and autoradiographed at −70°C with an intensifying screen.

Nonsynonymous and Synonymous Divergence Values

Nonsynonymous and synonymous divergence values were calculated using the method of Li (1993) and with a computer program kindly supplied by Wen-Hsiung Li.

Phylogenetic Analyses

Phylogenetic analyses were carried out using the program PAUP (version 3.1, Swofford 1993) as specified in the results section. Bootstrap analysis was also carried out using PAUP and a heuristic search option with 500 replicates.

Results

Southern Analysis of S. crassicaudata Genomic DNA: Evidence for a Two-Gene Cluster

Sequential Hybridization of Adult and Embryonic β-like Globin Probes

Southern analysis of DNA from animal 1650.1a using the S. crassicaudata ε-globin (S.c-ε) probe pSG-2H revealed strongest hybridization of the probe to 4.4-kb EcoRI, 4.5-kb BamHI, and 500-bp HindIII fragments (Fig. 1A). These fragments were shown previously

---

**Fig. 1.** Autoradiographs showing hybridization of (A) embryonic (pSG-2H) and (B) adult (pDG-5HS) β-globin probes to total genomic DNA from one animal (1650.1a). Enzymes used are: E: EcoRI, B: BamHI, and H: HindIII. Fragment lengths in kilobases (kb) were estimated by comparison with a λ-HindIII size marker.
(Cooper and Hope 1993) to contain a single embryonic β-like globin gene. When the filter was stripped and reprobed with pDG-SHS (D. viverrinus adult β-globin), strongest hybridization was detected between the probe and 4.0-kb and 3.2-kb EcoRI, 27-kb BamHI, and 15-kb HindIII fragments, which had also cross-hybridized to the S.c-ε probe (fig. 1). These results suggest the above fragments contain a putative adult β-globin gene. In addition, the S.c-ε probe hybridized less strongly to 7.6-kb EcoRI, 15-kb BamHI, and 20-kb HindIII fragments (fig. 1). Each of these fragments failed to hybridize to the adult β-globin probe.

Detection of BamHI and EcoRI Restriction Fragment Variation

Southern analysis of total genomic DNA from family M1920 (two parents, six offspring) digested with BamHI and hybridized with the S.c-ε probe pSG-2H, revealed restriction fragment variation involving three fragments of length 4.5 kb, 6.5 kb, and 8.9 kb (henceforth designated B1-4.5, B1-6.5, and B1-8.9, respectively). The inheritance of these fragments was in accord with Mendelian expectations for the inheritance of co-dominant allelic variants at an autosomal locus. One of these fragments (B1-4.5) was shown previously to contain the S. crassicaudata ε-globin gene (Cooper and Hope 1993).

Southern analysis of total genomic DNA from the family M1625 (two parents, eight offspring), digested with EcoRI and hybridized with the S.c-ε probe, revealed a complex array of between three and six fragments that hybridized to the probe in each individual (fig. 2A). The inheritance of these fragments can most simply be interpreted as follows: The 11.4-kb and 4.4-kb variants (E1-4.4 and E1-11.4) contain regions of the S. crassicaudata ε-globin gene (S.c-ε) and result from the presence or absence of an EcoRI site located approximately 3 kb upstream of S.c-ε (see Hope et al. 1992). The 4.0-kb, 3.4-kb, and 3.2-kb variants (E2-4.0, E2-3.4, and E2-3.2) contain regions of an adult β-globin gene (henceforth referred to as the S.c-β locus). The 7.6-kb and 6.8-kb variants (E3-7.6 and E3-6.8) contain regions of a third sequence (henceforth referred to as the ε2 locus). Southern analysis of DNA from 73 animals in the S. crassicaudata colony revealed examples of six genotypes of E2 variants at the S.c-β locus, three genotypes of E3 variants at the ε2 locus, and two genotypes of E1 variants at the S.c-ε locus (fig. 2B).

Linkage Analysis of the S.c-ε, S.c-β, and ε2 Loci

Parents of 14 litters were typed for BamHI variants at the S.c-ε locus, EcoRI variants at the S.c-ε, S.c-β, and ε2 loci. These typings were used to determine for each pair of loci which families contained one doubly heterozygous parent and one doubly homozygous parent. Five such families were detected and their offspring were typed for EcoRI and BamHI restriction site variation at each locus (data available on request). Linkage analysis of the S.c-ε and S.c-β loci revealed a maximum Likelihood of 0.01, providing strong evidence that the ε- and β-globin loci of S. crassicaudata are closely linked (table 1A). Linkage analysis of the S.c-ε and ε2 loci revealed a maximum Likelihood of 0.02 at a recombination fraction of 0.45, and z < -2 at a recombination fraction of 0.1 (table 1B). These data provide evidence that the
and β2 loci are not closely linked. Linkage analysis of the S.c-β and β2 loci revealed a Lod score of z < -2 at a recombination fraction of 0.1, providing evidence that S.c-β and β2 also are not closely linked (table 1C).

Isolation of an Adult β-Globin Gene

Screening of a genomic DNA library using the marsupial adult P-globin probe pDG-5R resulted in the isolation of two phage clones, one of which was analyzed in detail (designated XSG-6). Southern blot analysis of XSG-6 DNA, probed with pSG-2H(S.c-ε probe), revealed hybridization of the probe to a 3.5-kb EcoRI fragment, diagnostic of an adult P-like globin gene (see Southern data above). The genomic insert of λSG-6 was restriction mapped and found to overlap approximately 4.5 kb with the phage clone λSG-1, previously isolated by Cooper and Hope (1993) and shown to contain the third exon and 3' flanking region of S.c-ε (fig. 4). This result was confirmed by Southern analysis of XSG-6 DNA using an 860-bp EcoRI-SalI fragment derived from the terminal 3' end of XSG-1 as a probe. This probe was found to hybridize to 1.5-kb EcoRI, and 2.5-kb and 0.8-kb HindIII fragments (data not shown) which map to the 5' side of the phage clone ASG-6 (fig. 3).

In order to sequence the putative adult P-globin gene in ASG-6, the 3.5-kb EcoRI fragment, which had cross-hybridized to the S.c-ε probe pSG-2H, was subcloned into pGEM-3Zf(-) and hybridization to the S.c-ε probe further localized to a 700-bp NcoI/NcoI fragment and a contiguous 600-bp NcoI/HindIII fragment. These fragments were subcloned and sequenced from the M13 reverse and forward sequencing primers. In addition, a combination of primers was designed to sequence the entire gene on both strands using a PCR/cycle sequencing approach (fig. 4). This gene is henceforth referred to as S.c-β.

Structure of the Gene

The location of the putative cap site and 5' promoter sequences, including the “ATA box,” “CAAT box,” and two “CACCC boxes,” a 3' poly(A) signal (AATAAA), and conserved intron/exon splicing signals were identified by their homology with corresponding regions of the D. virginiana adult P-globin gene. Comparison of the promoter sequences with corresponding sequences in embryonic and adult P-globin genes from marsupials and eutherians indicates that promoter sequences of S.c-ε most closely resemble those found in the D. virginiana adult P-globin gene (data not shown). In addition, the presence of two CACCC boxes is a characteristic of adult P-globin genes of D. virginiana and eutherian mammals, while just one CACCC box is found adjacent to marsupial and eutherian embryonic P-like globin genes (Dierks et al. 1983; Myers, Tilly, and Maniatis 1986; Koop and Goodman 1988).

The two introns split the gene at positions identical to those in all known mammalian P-like globin genes. The first intron of S.c-β (112 bp) is similar in length to those found in D. virginiana and most eutherians which range from 109 bp to 132 bp in length. The second
**β-globin gene cluster of S. crassicaudata**

**FIG. 3**—Phage map of the β-globin gene cluster of S. crassicaudata showing the linkage order of the two genes S. c-ε and S. c-β. Three of the phage clones, XSG-5, XSG-1, and XSG-3, were previously reported by Cooper and Hope (1993). Enzymes mapped are EcoRI (E), HindIII (H), and BamHI (B), and sizes are in kilobases (kb). Also shown are the structures of each gene, with coding regions boxed, and lengths of the first and second introns in base pairs (bp). Regions of overlap of each phage clone were verified by Southern blot analysis. A 7.5-kb EcoRI fragment was unable to be mapped on the phage clone XSG-6, *E* refers to a Hind111 site which was present in the phage clone XSG-6 but absent in the overlapping clone XSG-1 and is most likely to represent polymorphic restriction site variation as both clones were isolated from independent genomic DNA libraries.

**FIG. 4**—Nucleotide and inferred amino acid sequence of the S. crassicaudata adult β-globin gene (S. c-β). The gene is divided into three exons and two introns flanked by conserved donor and acceptor splicing signals (underlined). Also underlined are 5′ promoter signals “CACCC” and “CCAAT,” an RNA polymerase II binding site “ATA,” a putative cap site, and a poly(A) signal “AATAAA.” The inferred amino acid sequence is 146 residues long and given in the standard three-letter code. Oligonucleotide primers used in the sequencing of the gene are highlighted in bold with reverse strand oligonucleotides also underlined. The lengths of the first and second introns are 112 bp and 1,541 bp, respectively. The entire gene was sequenced on both strands with the exception of the last 90 bp. GenBank accession number for the gene is 269592.
introns of \( S_c \)-\( C \) is 1,541 bp in length, similar to the second intron of \( S_c - 1 \) (1,438 bp) and the adult P-globin gene of \( D. \) virginiana (1,465 bp), and 16 bp shorter than the -globin second intron of \( D. \) virginiana. The coding region of \( S_c \)-\( C \) conceptually translates to give a polypeptide chain consisting of 146 amino acids. Comparison of the first 60 amino acids of this chain with the sequence derived from the P-chain of \( S. \) crassicaudata adult hemoglobin (Cooper and Hope 1993) shows just two differences out of 60 amino acids, at sites 10 (His-Val) and 13 (Ala-Gly). Given that the DNA sequence of \( S_c \)-\( C \) was determined from both strands, these differences are probably due to either polymorphic variation in the P-globin chain of \( S. \) crassicaudata or possible errors in the derivation of the polypeptide sequence reported by Cooper and Hope (1993). In the corresponding region of the \( S_c \)-\( C \) chain there are 20 amino acid differences with the adult P-chain. These results provide strong evidence that \( S_c \)-\( C \) is the adult P-globin gene of \( S. \) crassicaudata.

**Nonsynonymous and Synonymous Divergence Values**

At nonsynonymous and synonymous sites, \( S_c \)-\( C \) was found to most closely resemble the dasyurid native cat \( Dasyurus \) viverrinus partial adult \( \beta \)-globin sequence (7% and 3% divergence, respectively) and the \( D. \) virginiana adult \( \beta \)-globin gene (13% and 42% divergence, respectively; table 2). \( S_c \)-\( C \) differs from the \( D. \) virginiana and \( S. \) crassicaudata embryonic P-like globin genes by 17% and 15%, respectively, at nonsynonymous sites and by 109% and 99%, respectively, at synonymous sites (table 2). These data indicate that \( S_c \)-\( C \) is orthologous to marsupial adult \( \beta \)-globin genes. \( S_c \)-\( C \) was also found to differ at nonsynonymous sites from eutherian adult \( \beta \)-globin genes by between 21% and 23%. This contrasts with the low divergence between marsupial and eutherian embryonic \( \beta \)-like globin genes (9%–14%, Cooper and Hope 1993) and indicates that the rate of evolution of nonsynonymous sites in embryonic P-like globin genes in mammals is almost two thirds slower than the corresponding rate in adult \( \beta \)-globin genes (for further discussion see Hardison 1984; Koop and Goodman 1988; Cooper and Hope 1993).

**Phylogenetic Analyses**

A phylogenetic analysis of adult and embryonic P-like globin genes from marsupials and eutherians was carried out to investigate the evolution of the P-globin gene family in mammals. Only coding regions of each gene were used in the analysis and the chicken adult \( \beta \)-globin sequence was used as an outgroup. A parsimony analysis using the program PAUP (version 3.1, Swofford 1993) was carried out with a heuristic search option, weighting each site equally, and four most-parsimonious trees of length 1,000 were obtained. One of the trees depicts the embryonic and adult genes in separate monophyletic groups with each containing individual sister groups of marsupial and eutherian genes (fig. 5A). A strict consensus of the four minimal trees indicates that the branching pattern of the four eutherian \( \epsilon \)-globin genes and the relative position of the marsupial
embryonic clade were unresolved (fig. 5B). The individual sister groups containing marsupial and eutherian adult genes were supported by 99% and 95% of bootstrap replicates, respectively, and monophyletic groups of marsupial and eutherian embryonic genes were each supported by 79% and 40% of bootstrap replicates, respectively. Weighting first and second codon positions two-fold over third positions to account for the faster substitution rate at synonymous sites did not alter the major branching patterns observed above (data not shown). Each of the major monophyletic groups referred to above were also present in a phylogenetic tree constructed using the neighbor-joining distance method (in PHYLIP version 3.4; Felsenstein 1988; Saitou and Nei 1987; data not shown). These results provide evidence that the adult and embryonic β-like globin gene lineages arose by duplication of an ancestral β-globin gene prior to the separation of the marsupial and eutherian lineages.

Northern Analysis

Sequence comparisons of individual exons of S.c-ε and S.c-β revealed that the first exon was the most variable (33% divergence compared with 21% and 26% in the second and third exons, respectively). The first exon of S.c-ε was therefore PCR-amplified and used as an S.c-ε-specific probe of RNA isolated from an embryo and neonatal (henceforth referred to as pouch young) of S. crassicaudata to investigate the developmental timing of expression of S.c-ε. A 700-bp Nco I fragment in the plasmid subclone pSG-6, containing the first exon, first intron, and 97-bp of the 5’ end of the second exon of S.c-β, was used as a S.c-β-specific probe. Both probes were tested for their specificity by Southern analysis of ASG-3 (containing S.c-ε) and ASG-6 (containing S.c-β) DNA (data not shown). Appropriate washing conditions were then determined to ensure that cross-hybridization of the two probes would be reduced to a minimum. The S.c-ε-specific probe was found to hybridize to RNA isolated from an embryo and 1-day-old pouch young (fig. 6). The S.c-β-specific probe failed to hybridize to embryonic and 1-day-old pouch young RNA and showed weak hybridization with RNA isolated from pouch young between 2 and 6 days old (post partum) and strong hybridization to RNA isolated from a 36-day-old animal (fig. 6). When the filter was stripped and reprobed with an Aspergillus niger rDNA sequence (a gift of M. O’Connell, University of Adelaide) the strongest hybridization was detected between the probe and 2-, 4-, and 6-day-old pouch young RNA (data not shown). These results suggested that a higher level of globin mRNA should be present in these lanes than for the 36-day-old animal. However, a quantitative analysis of β-like globin RNA or rRNA in each lane could not be carried out, due to degradation of some of the RNA on the Northern (resulting in smear hybridization signals, see fig. 6). Although it is most likely that this weaker hybridization signal of the S.c-β probe to RNA isolated from pouch young less than 6 days old resulted from the degradation of RNA in these lanes, it cannot be ruled out that RNA from a third β-like globin gene may be present at this time (see Discussion for reasons why this latter possibility is unlikely). Taken overall, these results provide evidence that S.c-ε is expressed in embryonic tissues and switched off or reduced in expression within a day or two of birth, and that S.c-β is expressed in pouch young and adult tissues.

Discussion

Southern analyses of S. crassicaudata genomic DNA and molecular characterization of a lambda clone of an adult β-globin gene, together with data from a previous study of an embryonic β-like globin gene from S. crassicaudata (Cooper and Hope 1993), provide evidence that just two functional β-like globin loci, one embryonic (S.c-ε) and one adult β-globin gene (S.c-β), exist in this marsupial. These results also verify that the two genes are closely linked to each other in the order 5’-ε-β-3’ spanning a distance of approximately 26 kb. Southern and genetic analyses indicate that both genes are unlinked to a third locus (ε2) detected with the S.c-ε probe pSG-2H. The ε2 locus consisted of DNA fragments which hybridized to pSG-2H but failed to hy-
bridize to an adult β-globin probe pDG-5HS (see fig. 1). These data, and the location of the e2 locus away from the β-globin gene cluster, suggest that it is unlikely that e2 is a functional P-like globin gene. The e2 locus is therefore likely to contain either regions of a highly diverged β-globin pseudogene or an independently derived region of DNA which shares sequence similarity with a portion of the S. c-e gene (for example, a dispersed repetitive sequence in the first or second introns of S. c-e). Although just two P-like globin genes were detected in the phage clones isolated in this study, the possibility remains that further highly diverged β-like globin pseudogenes are also present in the cluster which could not be detected by Southern analysis.

The orientation of the two P-like globin genes of S. crassicaudata is analogous to the 5'-3' orientation of embryonic and adult sets of β-like globin genes in all eutherian β-globin gene clusters studied to date (see Goodman et al. 1984; Hardison 1991). These results suggest that it is likely that the eα-globin and βm-globin genes of the marsupial D. virginiana are also closely linked in a similar arrangement, as predicted by Koop and Goodman (1988). The presence of a two-gene cluster in representatives of each of the two major cohorts of marsupials (Ameridelphia and Australidelphia) (Aplin and Archer 1987), provides strong evidence that a two-gene ancestral cluster existed prior to the separation of these major marsupial lineages approximately 105 MYBP. In addition, the conserved 5'-3' linkage order of the adult and embryonic β-like globin genes of S. crassicaudata and all known eutherian β-globin gene clusters lends support to the theory that a two-gene ancestral cluster existed prior to the divergence of the eutherian and marsupial lineages, and that both genes were already differentiated with respect to their mode of developmental expression (Koop and Goodman 1988; Cooper and Hope 1993).

Further support for these conclusions comes from a phylogenetic analysis of β-like globin genes from marsupials and eutherians. A consensus of four minimal trees depicted each of the adult sequences forming a monophyletic group to the exclusion of the embryonic sequences, and marsupial and eutherian embryonic genes in separate monophyletic groups. This observation is consistent with the hypothesis that two P-like globin genes, including one embryonic- and one adult-expressed gene, were present in the most recent common ancestor of marsupials and eutherians. Among eutherian embryonic sequences there were two monophyletic groups containing y-like and q-like globin genes (fig. 5). The phylogenetic results are therefore in accord with the proposal that the proto-y and proto-γ-globin genes arose through tandem duplication of a ancestral e-globin gene after the divergence of the marsupial and eutherian lineages and prior to the separation of the lagomorph, rodent, artiodactyl, and primate lineages (Goodman et al. 1984).

Northern analyses of embryonic and pouch young total RNA reported in this paper support previous findings that the expression of S. c-e is switched off or reduced within a day or two of birth (Cooper and Hope 1993). There is also evidence to suggest that embryonic β-like globin mRNA was present until 1 day after birth, although in an independent study Cooper and Hope (1993) detected embryonic P-like globin mRNA in pouch young up to 4-days-old. These results were confirmed in a recent study by Holland et al. (1994) using the technique of isoelectric focusing of hemoglobin protein. Two hemoglobins were observed in newborn marsupials that disappeared from the blood by day 5. Concomitantly, the red cells at birth were all found to be nucleated (embryonic-like) but by day 4 only 15% of the cells were nucleated (Holland et al. 1994). Results from the present study suggest that adult β-globin mRNA is absent in embryonic tissues and therefore just a single β-like globin chain (S. c-e chain) would be found in hemoglobin prior to birth. The observation of two forms of embryonic hemoglobin would therefore imply that two different types of a-globin chain are present. Alternatively, a second embryonic form of hemoglobin may originate from the aggregation of two or more tetramers (Holland et al. 1994). In other marsupials such as the tammar wallaby Macropus eugenii and the brush-tail possum Trichosurus vulpecula, four distinct embryonic hemoglobins have been detected, which suggests that these species may have extra embryonic P-like or a-like globin chains or genes which are not found in S. crassicaudata (Hope 1970; Holland et al. 1988; Calvert, Holland, and Gemmell 1994).

The Northern analysis using an adult-specific probe showed a relatively weak hybridization signal with RNA from pouch young between 2 and 6 days old compared with RNA from a 36-day-old animal. It is most likely that this resulted from reduced levels of intact β-like globin gene, this would seem unlikely given all the other lines of evidence to suggest that just two functional P-like globin genes exist in S. crassicaudata. In addition, Holland et al. (1994) observed, using isoelectric focusing of S. crassicaudata pouch young hemoglobin, the first appearance of adult hemoglobin at day 3 and complete replacement of nonadult hemoglobin by day 5. Interestingly, they also observed different oxygenation properties of hemoglobin in pouch young between 6 and 12 days old (bent Hill plots of oxygen equilibrium curves) relative to those of adult animals (straight Hill plots), and concluded that this resulted from higher concentrations of 2,3-diphosphoglycerate in adult animals than in pouch young. Taking all the available evidence into account it is most likely that there is a single switch in β-globin gene expression from Sc-e to S.c-β within a day or two of birth.

This study has provided strong evidence for the existence of a two-gene β-globin cluster in S. crassicaudata, consisting of one embryonic- and one adult-expressed gene, linked in the order 5'-e-β-3' and spanning 26 kb. The simplicity of this arrangement and their pattern of expression provides for an ideal model system to test hypotheses and models concerning how the
β-globin gene family is coordinately and differentially expressed during development.

Acknowledgments

We thank H. Bennett and D. Golding for providing animals from the *S. crassicaudata* colony, S. Kolz and C. Chesson for help with DNA extractions, and B. Wainwright for providing a marsupial β-globin probe. We also thank two anonymous reviewers for their helpful comments on the manuscript. S.J.B.C. received support from a Commonwealth Postgraduate Research Award, the University of Adelaide, and the S.A. Museum.

LITERATURE CITED


THOMAS EICKBUSH, reviewing editor

Accepted May 14, 1996