Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 10757–10761, November 1995
Evolution

Emergence of the *ZNF91* Krüppel-associated box-containing zinc finger gene family in the last common ancestor of Anthropoidea

(zinc finger proteins/gene amplification/primate evolution/gene clusters)

ERIC J. BELLEFROID*†‡§, JEAN-CHRISTOPHE MARINE§¶‖, A. GREGORY MATERA*, CATHERINE BOURGUIGNON*†, TRUSHNA DESAI*, KIM COLEMAN HEALY*, PATRICIA BRAY-WARD*, JOSEPH A. MARTIAL‖, JAMES N. IHLE¶, AND DAVID C. WARD*

*Department of Genetics and Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510; ¶Department of Biochemistry, St. Jude Children's Research Hospital, 332 North Lauderdale, Memphis, TN 38105; and ‖Laboratoire de Biologie Moleculaire et de Génie Génétique, Université de Liège, Institut de Chimie B6, B-4000 Sart-Tilman, Belgium

Communicated by Edward A. Adelberg, New Haven, CT, May 19, 1995

ABSTRACT **The *ZNF91* gene family, a subset of the Krüppel-associated box (KRAB)-containing group of zinc finger genes, comprises more than 40 loci; most reside on human chromosome 19p12–p13.1. We have examined the emergence and evolutionary conservation of the *ZNF91* family. *ZNF91* family members were detected in all species of great apes, gibbons, Old World monkeys, and New World monkeys examined but were not found in prosimians or rodents. In each species containing the *ZNF91* family, the genes were clustered at one major site, on the chromosome(s) syntenic to human chromosome 19. To identify a putative "founder" gene, >20 murine KRAB-containing zinc finger protein (ZFP) cDNAs were randomly cloned, but none showed sequence similarity to the *ZNF91* genes. These observations suggest that the *ZNF91* gene cluster is a derived character specific to Anthropoidea, resulting from a duplication and amplification event some 55 million years ago in the common ancestor of simians. Although the *ZNF91* gene cluster is present in all simian species, the sequences of the human *ZNF91* gene that confer DNA-binding specificity were conserved only in great apes, suggesting that there is not a high selective pressure to maintain the DNA targets of these proteins during evolution.**

The zinc finger nucleic-acid binding motif (Cys_2–His_2) occurs in hundreds to thousands of vertebrate proteins and is traceable throughout eukaryotes (1). Although structural features common to various subsets, including the 7-amino acid "H/C link" between adjacent fingers (2, 3), and various conserved N-terminal modules (4–8) have been described, functions are known only for the few zinc finger proteins (ZFPs) that act as transcription regulators in differentiation and development; the functions of most mammalian ZFPs are unknown (refs. 9–15; reviewed in ref. 16).

Several related ZFP genes in human and *Xenopus* are arranged in clusters (17–21). In humans, the ZNF91 family, containing >40 highly related Krüppel-associated box (KRAB)-containing ZFP genes or pseudogenes (KRAB–ZFP), has recently been identified on chromosome 19p12–p13.1 (22). *ZNF91* family members share a common exon–intron organization, with all zinc finger repeats in a single exon and the KRAB domain encoded by two separate exons. The linker region between the N-terminal KRAB domain and the first finger is conserved in length and sequence. ZNF91 family members are transcribed in many tissues but most abundantly in T-lymphoid cells.

Here we have studied the emergence and evolutionary conservation of the *ZNF91* family in search of clues to its function. We sought homologous genes in rodents and in a panel of primate species. We also attempted to clone orthologs of the *ZNF91* gene by PCR with primers corresponding to the zinc finger amino acids determining DNA-binding specificity.**

MATERIALS AND METHODS

Genomic DNAs. Genomic DNAs were isolated as described (23) from the following cell lines or tissue samples: human genomic DNA from HeLa cells; *Pan troglodytes* (chimpanzee), *Gorilla gorilla* (gorilla), and *Pongo pygmaeus* (orangutan) DNA from lymphoblastoid cell lines provided by J. Kidd (New Haven, CT), *Hylobates syndactylus* (gibbon) DNA from the lymphoblastoid cell line provided by J. Wienberg (Munich); genomic DNAs from *Lemur fulvus fulvus* (brown lemur), *Propithecus verreauxi* (sifaka), *Loris tardigradus* (slender loris), *Galago moholi* (lesser bushbaby), and *Tarsius syrichta* (Philippine tarsier) from tissue samples from Duke University Primate Center, (Durham, NC); genomic DNAs from *Nasalis larvalus* (proboscis monkey), *Presbytis entellus* (Indian langur), and *Cebus apella* (capuchin) from peripheral blood lymphocyte samples provided by J. Powell (New Haven, CT). *Symphalangus syndactylus* (siamang), *Theropithecus gelada* (gelada baboon), *Colobus polykomos* (colobus monkey), *Erythrocebus patas* (patas monkey), and *Aotus trivirgatus* (owl monkey) DNAs were a gift from S. J. O'Brien (National Institutes of Health, Bethesda). *Mandrillus sphinx* (mandrill) and *Saimiri boliviensis* (squirrel monkey) DNAs were a gift from J. Rodgers (San Antonio, TX). *Callithrix jacchus* (marmoset) DNA was a gift from G. Miller (New Haven, CT).

RNA Isolation and Reverse Transcription–PCR. Total RNA was extracted by the guanidinium isothiocyanate method (24) and was purified by centrifugation through cesium chloride. Single-stranded cDNA was prepared with oligo(dT) primers and the cDNA cycle kit (Invitrogen). Twenty-five nanograms of the cDNA mixture was diluted in 50 μl of PCR mixture and amplified as described below.

PCR. PCR primers and cycling conditions are described in the figure legends. PCRs were performed in 100-μl reaction mixtures containing 0.2–1 μg of genomic DNA, 1.5 mM $MgCl_2$, 0.2 mM dNTP, 0.5 μM primers, and 2.5 units of *Taq* poly-

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.

Abbreviations: ZFP, zinc finger protein; KRAB, Krüppel-associated box; YAC, yeast artificial chromosome.

†Present address: Georg-August-Universität Göttingen, Institut für Biochemie und Molekulare Zellbiologie, Humboldtalle 23, D-37073, Göttingen, Germany.
‡To whom reprint requests should be addressed.
§E.J.B. and J.-C.M. contributed equally to this work.

**The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U29495–29514 and U29524–U29527).

merase (Perkin–Elmer/Cetus or Promega) using the supplier's buffer.

Cloning and Sequencing of PCR Products. PCR products were made blunt-ended with Klenow enzyme and cloned into the pGEM3 (Promega) vector. Alternatively, PCR products were cloned in the TA vector from Invitrogen. Inserts were sequenced with the Sequenase kit (United States Biochemical). Sequence analysis utilized the software package of the University of Wisconsin Genetics Computer Group (Madison, WI).

Nucleic Acid Hybridization. Southern blots were done exactly as described using Biotrans membrane (ICN) (22). Final washes were at 50°C in 2× SSC/0.1% SDS. cDNA probes were ^{32}P-labeled using the Prime-It kit (Stratagene). Oligonucleotides were ^{32}P-end-labeled with [^{32}P]ATP and T4 polynucleotide kinase. Membranes were exposed to Kodak XRP-5 films at −70°C for 2–24 h.

***In Situ* Hybridization.** Metaphase chromosome spreads were prepared by standard methods (25). Human chromosomes were obtained from peripheral blood lymphocytes; great ape and gibbon chromosomes were from lymphoblastoid cell lines; *Tarsius* and *Galago* chromosomes were from fibroblasts cultured from biopsies obtained from the Duke University Primate Center (Durham, NC). The chromosome-specific library pBS19 (27) and DNA from the yeast artificial chromosome (YAC) clones 71D9, 179B7, 245F12, 143B11, 414H7, 60D10, 138D1, 332C9, 158C9, 335H4, 84A10, 131F11, 88C4, 96A1, 140G4, 143B6, 143D6, and 16B10 of the CEPH library, encompassing the ZNF91 KRAB–ZFP gene cluster on chromosome 19p12–p13.1, was labeled with either biotin or digoxigenin by standard nick-translation methods. *In situ* suppression hybridization, detection, counterstaining, microscopy, image processing, and photography were done as described (22). The programs used for image capture (CCD IMAGE CAPTURE) and merging and pseudocoloring (GENE JOIN MAX PIX), written by T. Rand, are available for a fee from the Office of Cooperative Research at Yale University.

RESULTS

Appearance of KRAB–ZFP Genes of the *ZNF91* Family in Higher Primates. The *ZNF91* KRAB–ZFP genes have been identified in humans but are undetectable by hybridization in mouse, rat, and hamster genomic DNAs, suggesting that the *ZNF91* gene family evolved later than Rodentia (22). To further investigate the emergence time of this KRAB gene subfamily, we screened a panel of primate genomic DNAs by PCR. Since the sequence of the link region between the KRAB box and the first finger domain is specific to KRAB–ZFP genes of the *ZNF91* family, we designed degenerate primers to the most conserved regions of this *ZNF91*-specific element. PCR products were electrophoresed, blotted, and hybridized with a ^{32}P-labeled *ZNF91* cDNA spacer probe to confirm the identity of the amplified DNA. PCR products of the expected size (250 bp) were observed in all ape, New World monkey, and Old World monkey samples (Fig. 1), but none was detected in prosimian species (tarsier, loris, galago, lemur, and sifaka). To test the efficiency of the PCR amplifications, we used degenerate primers directed against the more widely conserved KRAB domain and hybridized the PCR products with a *ZNF91* KRAB oligonucleotide probe. The expected 91-bp PCR products were observed in all species examined, including prosimians and rodents. We estimated the number of *ZNF91*-related sequences by slot-blot analysis of primates' genomic DNA using the *ZNF91* spacer probe, with a c-myc probe as loading control (data not shown). ZNF-related genes in all simians show copy numbers similar to humans' (Fig. 1*B*): 25–35 per haploid genome (as estimated by comparing signal intensity to serial dilutions of *ZNF91* DNA).

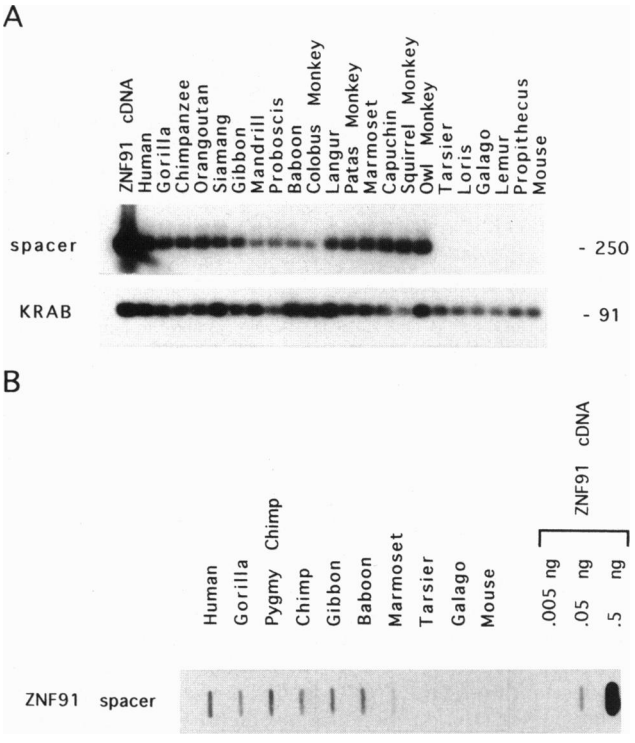

FIG. 1. Detection by PCR and dot blot analysis of *ZNF91* homologous genes in primates. (*A*) PCR products separated by electrophoresis and analyzed by Southern blotting. PCR conditions: 94°C for 3 min, followed by 35 cycles of 1 min at 94°C, 1 min at 50°C, and 1 min at 72°C. The last cycle was followed by a final elongation step (72°C for 10 min). Coordinates (22) and sequences of the degenerate primers: (*Upper*) nt 359–386, 5′-ATT TT(G/C) C(C/T)C (A/G)AG A(T/C)(T/C) TTT G(G/T)C C(A/G)G AGC-3′; nt 585–609, 5′-TGT CT(G/A) TTT GAA TTT (G/A)AA A(T/A)T (T/A)TA T-3′; nt 215–234, 5′-(C/T)T(G/A) (T/A)AG TTC TC(C/T) A(G/A)C AT(C/T) AC-3′. The KRAB and spacer probe were *ZNF91* cDNA, nt 168–188 and 359–609, respectively. (*B*) Dot blot analysis of 2 μg of genomic DNA from each species and serial dilutions of *ZNF91* cDNA using a *ZNF91* spacer probe (nt 372–693). Membranes were exposed to autoradiographic films for 4 h.

We explored the chromosomal organization of the primate *ZNF91* genes by double-label fluorescence *in situ* hybridization, hybridizing simultaneously a human chromosome 19-specific library (to "paint" regions syntenic to human chromosome 19), and a pool of YAC clones containing the human *ZNF91* cluster from chromosome 19p12–p13.1 (Fig. 2). In human, orangutan, and chimpanzee, the chromosome 19 and the YAC probes cohybridized and labeled only one chromosome (PTR20 and PPY20) (28). In gibbon (*H. syndactylus*), the human chromosome 19 library painted two chromosome pairs, HL13 and HL17 (26), and the YAC probe gave a unique signal on chromosome HL13. In galago (*Galago crassicaudatus*) and tarsier (*T. syrichta*), the human chromosome 19 library painted part of one, as yet unidentified, chromosome, but the 19 synteny region had no *ZNF91* signal, in agreement with our PCR and hybridization analysis (see above). However, multiple *ZNF91* signals were detected on the telomeric regions of other chromosomes; these might represent cross-hybridization with other ZFP-enriched loci or other conserved sequences represented within the pool of YAC DNA.

In humans, the *ZNF91* gene family represents about one-third of the KRAB–ZFP genes and is expressed predominantly in T-lymphoid cells (22). To identify the founder gene(s) of the ZNF91 family, which might have had insufficient sequence similarity in the spacer region to be detected, we performed PCR on cDNA from a mouse T-cell line (DA-2), using primers in the conserved KRAB domain (element A) and in the H/C

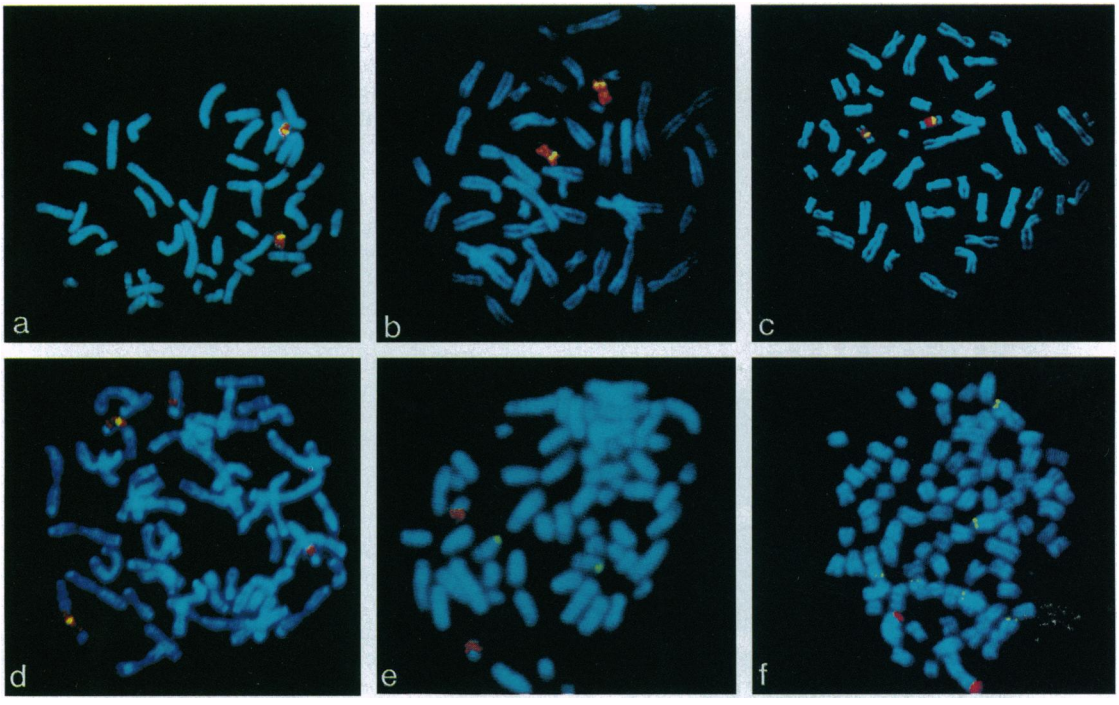

FIG. 2. Metaphase chromosomes of human (*a*), orangutan (*b*), chimpanzee (*c*), gibbon (*d*), galago (*e*), and tarsier (*f*) hybridized with a digoxigenin-labeled human chromosome 19-specific library probe (red) and biotin-labeled YAC DNAs derived from the human 19p12–p13.1 region including the *ZNF91* multigene cluster (yellow). Biotinylated probes were detected with fluorescein isothiocyanate-conjugated avidin; digoxigenin-labeled probes were detected with rhodamine-conjugated antibodies. Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole.

link fragments (29). The PCR products ranged from 200 to 2000 bp and represented the spacer regions and the first finger units of the KRAB–ZFP genes transcribed in this cell line. Sequencing of 50 individual clones detected 20 different spacer sequences (Fig. 3), indicating that the mouse, like the human, possesses a large number of KRAB–ZFP genes. *Zfp61* was the one most frequently found, being present in 23% of the clones. Also frequently detected were Zfp63 (14%), Zfp68 (9%), Zfp66 and Zfp70 (7%), and Zfp79 and Zfp80 (5%). The other cDNA clones were detected only once. None of the sequences shows any similarity to the human *ZNF91* spacer region or to other previously described ZFP genes. *Zfp10* to *-14* and *Zfp20* sequences show no KRAB-B homologous regions, which suggests that mouse KRAB–ZFP genes might have an exon–intron organization similar to that of the human genes and undergo alternative splicing (22).

***ZNF91* Sequences That Confer DNA-Binding Specificity Are Conserved Only in Great Apes.** Having demonstrated members of the *ZNF91* family in higher primates, we used PCR to search for true orthologs of the *ZNF91* gene and so estimate the degree of evolutionary conservation of individual family members. We tested our PCR analysis with primers to the first and third finger units of the human *EGR2* gene, which is highly conserved in vertebrates (30). As expected, PCR products were detected in all species (Fig. 4). The first set of *ZNF91*-specific primers tested lie at the 5′ and 3′ ends of the spacer and of the non-consensus type first finger element (primers a–b); they exclusively amplify the *ZNF91* gene from human DNA (22). They produced PCR products of the expected size in all simian primates (apes, Old World monkeys, and New World monkeys) (data not shown). PCR products containing the spacer sequence of the *ZNF91* orthologs from chimpanzee, orangutan, baboon, and capuchin were subcloned and sequenced. The amino acid sequences predicted from the chimpanzee and orangutan DNA sequences are identical to the human sequence. In contrast, the putative baboon and capu-

	KRAB-A	KRAB-B	SPACER	FINGER 1
ZNF91	YRNVMLENYRNLAFL	GIALSKPDLITYLEQGKEPWNMKQHEWVDEPT	GICPHFPQDFWPEQSMEDSFQKVLLRKYEKCGHENLQLRKGCKSVDECKVHKEGYNKLNQCLTTAQSKVFQCGKYLKVFYKFLNSNRHTIRH	TGKKCFKCKKCVKSFCIRLEKTQHKCVY
Zfp61	YRDVMFENYQNLATV	GYQLIKPSLISWLEQEEPSKQQKIVFPEWKLQ	LETKCSAFQQEFLRGNISSRMQMQTGINRGRELCDGTQYGDFFSELSPLRTDMRTQPAQDNTGCSQYRKDFLMLQRENCAGEELSEFSQ	
Zfp62	YREVMFENYRNVVSLA	GCFISKPDVISLLEQGKEPWKVVRQREPCPDS	ETKDETNQVFSENGIYENNLSQWKIMERVGNSGLKLLLKNGWESRREQERQEGSQGGISQSSEAHI	
Zfp63	YQDVMFENYRNLVSV	GYQLSKPTVISQLEKGEQPCMASSQGPEDPIL	LDVKNKLETKESTARDDISUKLDHGITRGRLIEDDIVCSPLKKASSYSDTLESSRATCGKGTRRAILDPUKETPRRG	
Zfp64	YREVMLENYNELVSL	GYQVSKPEVISNWSKEKSHGYARKESKDLSVQ	TGRPGLSPHGVLSRAYLKYSSEQMLSETTIGDTNNVAIQCHQESGRIHLGPKASSQ	
Zfp65	YREVMLENHSSVAGLA	GFLVFKPELIFRLEQGQEPWVLDLQGADGTEA	PRICQTDSAIRTDRKQTCETTSLLQRQIPGFGDNLDSKVWSEMLPRL	
Zfp66	YRDVMFENYNNLITV	GPPLTKPEVISKLEQEEEPCVVERFVLMRPCP	GEILGIDEHQKIQDGQVFEGIVVTSEASSCPEEFASTFFPNADSIPSMLSLFECDRCLKVF	
Zfp67	YREVMLENYSNLSFL	GYQVGEPRLISHLEQEEDLRWEERGIQQGPSP	VFSRLGTTSKEVEWSILLEDIPGHEASEDDVRLETSSLLETSSPSHLSISDEGPHLAQPVGRPTQKRALSPPQLNSQEQGSLDSA	
Zfp68	YRDVMLENYSNLLFL	GHCKTKPELIFNLEQRLGSWIVEASHQCVPDF	QNWSTLNKSSPDTKARHVWQVALNSETSKELAGIGNTCNVSTNHISDQTIKNENSLGMSPRKFISVEGTCISTLML	
Zfp69	YRDVMLENYSNFVSL	GLAISKPSIISLLEQGEPWIVEQAGWCPDLLS	VNAAKSLSKNGIPGKESLQKIIERVGSSCLESSCPGDDWKCKGLLSTKRGIYRGNIYRSDRPVGKEPSPRFRSAFPSAL	
Zfp70	YRNVMFENYQNLATV	GGQMFKPSLISWLEKKVELTVIRQGILQEWEN	HLKTKRTALQQDRFWSKDMSNGMQLWFPPEKC----------	TPEKSVECSDCGKTFVNQLELQTHSSSH
Zfp71	YREVMLENYSNLVFL	GLAFSKPYLVTFLEQSSDPWSMKRKESAAIH-	----------	LGSKPHKCKECGKAFDRNSVLIQHQRIH
Zfp72	YRDVMFENYSHLVFL	GLASCKPYLVTFLEQRQEPSVVERPAAATVHP	GGKPYKCQECFDKTFDWNSLLLQRQVH----------	FVEKPYKCEECGKVFSAHSALSIHQRIH
Zfp73	YREVMLENYSHLVFL	GLALSKPYLVTCLEKIQETSGVKRGAAPFMLP	GGKCTTCTECGKIFEWKEVFQNHQMLDD----------	LRMNPYKCEECAKSFRSPSLSSKHSRIH
Zfp74	YREVMFENYSNLVSL	GLAFSKPYLVTFLEKRQQSWDVKQQATVATTP	GTKPYKTKECNKTFVYRSLLIQHQRVH----------	TGEKPYKCEQCGKAFSCSSSLTPHQRIH
Zfp75	YREVMLENYSNLVTV	----------	ENTGICNTVCQHVVKTKKSCNELGEMLHEPSNCALTNRSDTTRASNNTCCKDQDASLDSSNFDRLKTHI	
Zfp76	YREVMLENYSNLVSV	----------	ENTCICNTVHQPVRTEKRSCNELNEMLHEPSNCALTNRSDTTENSNKTRCCNEDASVDSSAADRHKSNHTGEDTCQSKDCEKSFNLCSNI	
Zfp77	YREVMFENYRNLVSV	----------	GGKSVQETHELSLRQLVFEDVAVENTTSLGCESLTSGRNEKSKDLLLDRAMSQEIFNLKETVTHKGTNTKEIGYKHTESLKSAHLDSVVENIY	
Zfp78	YREVMLEKKLYACKQ	----------	CGKVFGRSSTLRKHERIH----------	TGEKPYLCKHCGKAFSDPTTRNNERTH
Zfp79	YREVMLENEQEKHMQ	----------	IVLQGNATRVTPKSQQSGKKAFWCKYDGCGKLTTTAHNLKVHERSH----------	TGDRPYQCEHSGKAFATGTGLKSHFRTH
Zfp80	YQDVMFENSTKDAVQ	----------	ERKEPQADVGYRPDDQRDHLSSKLIRRQSVP----------	TGENRYPCEHCGKAFEWALSLNQHKLSH

FIG. 3. Alignment of deduced amino acid sequence of KRAB–ZFP transcripts obtained by reverse transcription–PCR analysis of mouse lymphoma DA-2 mRNA with respect to the human ZNF91 sequence. Dashes denote deletions. Spacer region of Zfp61 to -69 and -70 to -77 was only partially sequenced. PCR conditions were the same as in Fig. 2 except for annealing (1 min 30 sec at 56°C) and elongation (2 min at 72°C). Coordinates (22) and sequences of the degenerate primers: nt 205–227, 5′-TA(T/C)(C/A) (G/A)(G/A)(G/A) A(C/G/T)G T(C/G)A TG(C/T) T(C/A)G AGA A-3′ (KRAB primer) and 5′-(A/G)(T/A)A (A/G/T/C)GG (T/C)TT (T/C)TC (A/G/T/C)CC (A/G/T/C)GT (A/G)TG-3′ (H/C link primer).

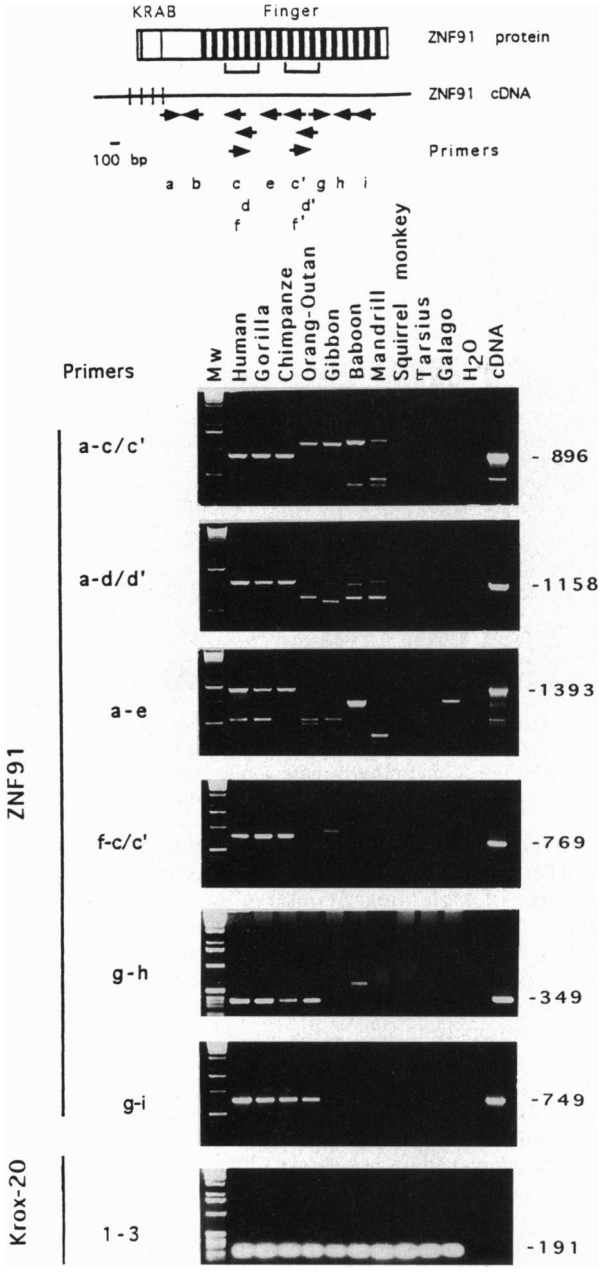

FIG. 4. PCR analysis of orthologous *ZNF91* genes in primates. Electrophoretic separation of PCR products: PCR conditions were the same as in Fig. 2, except for primer pairs a–c/c′ (annealing, 58°C; elongation, 1 min), a–d/d′ (annealing, 56°C; elongation, 1 min 30 sec), a–e (annealing, 52°C; elongation, 1 min), g–i (annealing, 56°C; elongation, 1 min). Coordinates and sequences (22) of the *ZNF91* primers: A (nt 373–396), 5′-CTTT TGG CCA GAG CAG AGC ATG TA-3′; B (nt 666–692), 5′-GTT GGG TTT TGT GTA AAC GGA TGC AA-3′; c (nt 1244–1269 and 2168–2193), 5′-GTT GAG AGT CGC TTA AAA GTT TTG TC-3′; D (nt 1508–1531 and 2432–2455), 5′-TCT AGT TAG GGT TGA AGA CCA TAT-3′; E (nt 1743–1766), 5′-TGT AGA AAG ACT TGA GGA ATG ATT-3′; F (nt 1424–1447 and 2248–2271), 5′-AAC TGG TCC TCA AGC CTT ACT AAA-3′; G (nt 2526–2539), 5′-AGC CGT TCC TCA ACC CTT ACT AAG-3′; H (nt 2852–2875), 5′-TGT AGT AAG GTG TGA AGG CTG GCT-3′; I (nt 3272–3275), 5′-TCC ATT TAG GGT TGA GGA TGA TAT-3′. Coordinates and sequences (30) of the human *EGR2* (Krox-20) primers: nt 1107–1130, 5′-TCC CGC TCT GAC GAG CTG ACA CGG-3′ (finger 1); nt 1275–1298, 5′-GCG CTT CCT CTC ATC ACT CCG GGC-3′ (finger 3).

chin orthologs encode proteins with only 75–78% sequence identity (Fig. 5).

Since the best criterion for orthology among ZFPs may be the conservation of the amino acids preceding the histidine residues conferring DNA-binding specificity (31), we next chose primers from the regions preceding the histidine residues of different units of the *ZNF91* finger domain (primers b–i, not primer a). Primers c, d, and f are present twice in the *ZNF91* cDNA sequence because it contains two groups of nearly identical tandemly repeated finger units (fingers 6–12 and 17–23). These primer pairs amplified products of the expected size only in humans and great apes (Fig. 4). With primer pairs a–c/c′, a–e, and f–c/c′, PCR products of the expected size could not be detected in orangutan DNA. With the primer pairs used for the *ZNF85* gene, another member of the *ZNF91* family, PCR products were also detected only in human and great apes (data not shown). This suggests that the conservation of the DNA-binding specificity of *ZNF91* genes is not under a high selective pressure.

DISCUSSION

The human *ZNF91* KRAB–ZFP family contains ≈40 highly related genes, most of which are clustered on chromosome region 19p12–p13.1 and are highly expressed in T-lymphoid cells (22). Our present studies reveal that the *ZNF91* family emerged after separation of the lineage Anthropoidea and that in the species containing the *ZNF91* family, the *ZNF91*-related genes cluster at one major site on the chromosomes syntenic to human chromosome 19. The presence and apparently similar copy number of the *ZNF91* gene family in all simian primates suggest that the duplication events producing this gene family occurred early in simian primate evolution, before the catarrhine primates (human, apes, and Old World monkeys) diverged from platyrrhines (New World monkeys). This correlates with the nucleotide divergence (26/258 nt) observed between the spacer-region sequences of the putative *ZNF91* ortholog genes from human and capuchin, the most distantly related species in which orthology has been identified. Using a rate of nucleotide substitution of 1.56×10^{-9} per site per year (32), the *ZNF91* gene is estimated to have appeared some 64 million years ago. *Tarsius*, whose relationship to other primates has been much disputed (33, 34), has no detectable *ZNF91* genes. The apparent absence of the *ZNF91* subfamily of KRAB–ZFP genes in rodents and strepsirhine prosimians (lemur, loris, galago, and sifaka) suggests that the mammalian common ancestor lacked this subfamily; its presence in anthropoids but not in *Tarsius* suggests that it arose in the most recent common ancestor of Anthropoidea.

The *ZNF91* family, with its recent origin, contrasts sharply with the conserved ZFPs that serve as developmental or differentiation-regulating transcription factors (16, 35–37). These ZFPs tightly conserve their finger domains (particularly the residues that determine binding specificity), whereas *ZNF91* finger domains are poorly conserved. PCR primers from DNA-binding domains failed to identify *ZNF91* orthologs in prosimians, Old World monkeys, or New World monkeys. Since *ZNF91* DNA-binding domains diverge even among primates, their target sequences may also be divergent. In both *Drosophila* (38) and *Xenopus* (39), ZFPs that interact with rapidly evolving, species-specific repetitive or transposable nucleic acid sequences have been described.

Genes encoding proteins important in the control of embryonic development are highly conserved, as exemplified by the homeobox genes (40). The recent evolutionary appearance of the *ZNF91* family suggests that its ZFPs do not regulate fundamental developmental processes. What then might be the function of genes of the *ZNF91* family? Although many possibilities exist, it is interesting to note that several ZFPs [i.e., su(Hw), Suvar] may participate in formation of chromatin structure (41). An examination of the intracellular distribution

```
Human KVLLRKYEKCGHENLQLRKGCKSVDECKVHKEGYNKLNQCLTTAQSKVFQCGKYLKVFYKFLNSNRHTIRHTGKKCFKCKKCVKSF
Chimpanzee ----------------------------------------------------------------------------------
O.Outan ----------------------------------------------------------------------------------
Baboon --I--RF--Y-P--------Y--------N-----AY---F---------SD-----LF--------K----E--SS-G-----P-
Capuchin -----R-K----------YR-------N-----GH---F--------RSE------F--------K----E--SS-G----FP-
```

FIG. 5. Sequence alignment of the spacer fragments obtained from the *ZNF91* orthologous genes of different primates. Dashes represent amino acid residues identical to the human ZNF91 sequence.

of the ZNF91 family proteins across the cell cycle might illuminate their potential biological function(s).

We are grateful to Drs. S. J. O'Brien, J. Kidd, G. Miller, J. Powell, J. Rodgers, and J. Wienberg for materials used in this study and to Dr. T. Haaf for helpful advice on chromosome preparations. We also thank Jon Lieman and Joan Menninger for helpful technical assistance and Dr. T. Pieler for comments on the manuscript. This work was supported in part by Grant HG-00272 (D.C.W.), the Services de Programmation de la Politique Scientifique—PAIP3-042, the Fonds National de la Recherche Scientifique—Télévie 7.4544.93, and the Caisse Générale d'Epargne et de Retraite-Financement de la Recherche Cancérologique 1993–1996 (J.A.M.) and by Support Grant P30 CA21765 from the National Cancer Institute Center (CORE), Grant CA51020 from the National Cancer Institute, and by the American Lebanese Syrian Associated Charities (ALSAC) (J.N.I.). J.-C.M. is a fellow of the Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture (IRSIA, Belgium).

1. Bellefroid, E. J., Lecocq, P. J., Benhida, A., Poncelet, D. A., Belayew, A. & Martial, J. A. (1989) *DNA* **8,** 377–387.
2. Rosenberg, U. B., Schröder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I. & Jäckle, H. (1986) *Nature (London)* **319,** 336–339.
3. Schuh, R., Aicher, W., Gaul, J., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kemler, R. & Jackle, H. (1986) *Cell* **47,** 1025–1032.
4. Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O. & Martial, J. A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3608–3612.
5. Rosati, M., Marino, M., Franze, A., Tramontano, A. & Grimaldi, G. (1991) *Nucleic Acids Res.* **19,** 5661–5667.
6. Knöchel, W., Pöting, A., Koster, M., El-Baradi, T., Nietfeld, W., Bouwmeester, T. & Pieler, T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6097–6100.
7. Klocke, B., Köster, M., Hille, S., Bouwmeester, T., Böhm, S., Pieler, T. & Knöchel, W. (1993) *Nucleic Acids Res.* **21,** 3767–3775.
8. Numoto, M., Niwa, O., Kaplan, J., Wong, K. K., Merrel, K., Yanagihara, K. & Calame, K. (1993) *Nucleic Acids Res.* **21,** 3767–3775.
9. Schneider-Maunoury, S., Topilko, P., Seitanidou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C. & Charnay, P. (1993) *Cell* **75,** 1199–1214.
10. Kinzler, K. W., Ruppert, J. M., Bigner, S. H. & Vogelstein, B. (1988) *Nature (London)* **332,** 371–374.
11. Schimmang, T., Lemaistre, M., Vorkamp, A. & Rüther, U. (1992) *Development (Cambridge, U.K.)* **116,** 799–804.
12. Kreidberg, J., Sariola, H., Loring, J., Maeda, M., Pelletier, J., Housman, D. & Jaenisch, R. (1993) *Cell* **74,** 679–691.
13. Morishita, K., Parker, D. S., Mucenski, M. L., Jenkins, N. A., Copeland, N. G. & Ihle, J. N. (1988) *Cell* **54,** 831–840.
14. Kreider, B. L., Orkin, S. T. & Ihle, J. N. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6454–6458.
15. Wimmer, E. A., Jäckle, H., Pfiefe, C. & Cohen, S. M. (1993) *Nature (London)* **366,** 690–694.
16. Pieler, T. & Bellefroid, E. J. (1994) *Mol. Biol. Rep.* **20,** 1–8.
17. Huebner, K., Druck, T., Croce, C. M. & Thiesen, H.-J. (1991) *Am. J. Hum. Genet.* **48,** 726–740.
18. Hoovers, J. M. N., Mannens, M., John, R., Bliek, J., Van Heyningen, V., Porteous, D. J., Leschot, N. J., Westerveld, A. & Little, P. F. R. (1992) *Genomics* **12,** 254–263.
19. Lichter, P., Bray, P., Ried, T., Dawid, I. B. & Ward, D. C. (1992) *Genomics* **13,** 999–1007.
20. Rousseau-Merck, M. F., Tunnacliffe, A., Berger, R., Ponder, B. A. J. & Thiesen, H. J. (1992) *Genomics* **13,** 845–848.
21. Nietfeld, W., Conrad, S., van Wijk, I., Giltay, R., Bouwmeester, T., Knöchel, W. & Pieler, T. (1993) *J. Mol. Biol.* **230,** 400–412.
22. Bellefroid, E. J., Marine, J.-C., Ried, T., Lecocq, P. J., Rivière, M., Amemiya, C., Poncelet, D. A., Coulie, P. G., de Jong, P., Szpirer, C., Ward, D. C. & Martial, J. A. (1993) *EMBO J.* **12,** 1363–1374.
23. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
24. Chirgwin, J., Przybyla, R., MacDonald, R. & Rutter, W. (1979) *Biochemistry* **18,** 5294–5299.
25. Lichter, P., Tang, C. C., Call, K., Hermanson, G., Evans, G. A., Housman, D. & Ward, D. C. (1990) *Science* **247,** 64–69.
26. Koehler, U., Arnold, N., Weinberg, J., Tofanelli, S. & Stanyon, R. (1995) *Am. J. Phys. Anthropol.* **97,** 37–47.
27. Collins, C., Kuo, W., Seagraves, R., Fuscoe, J., Pinkel, D. & Gray, J. W. (1991) *Genomics* **11,** 997–1006.
28. Jauch, A., Weinberg, J., Stanyon, R., Arnold, N., Tofanelli, S., Ishida, T. & Cremer, T. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8611–8615.
29. Ihle, J. N., Smith-White, B., Sisson, B., Parker, D., Blair, D. G., Schultz, A., Kozak, C., Lunsford, R. D., Adkew, D., Weinstein, Y. & Isfort, R. J. (1989) *J. Virol.* **63,** 2959–2966.
30. Joseph, L. J., Le Beau, M. M., Jamieson, G. A., Acharya, S., Shows, T. B., Rowly, J. D. & Sukhatme, V. P. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7164–7168.
31. Sommer, R. J., Retzlaff, M., Goerlich, K., Saunder, K. & Tautz, D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10782–10786.
32. Sakoyama, Y., Hong, K.-J., Byun, S., Hisajima, H., Ueda, S., Yaoita, Y., Hayashida, H., Miyata, T. & Honjo, T. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1080–1084.
33. Dutrillaux, B. (1979) *Ann. Genet.* **48,** 251–314.
34. Healy, K. C. (1992) Ph.D. dissertation (Duke Univ., Durham, NC).
35. Wimmer, E. A., Jackle, H., Pfeifle, C. & Cohen, S. (1993) *Nature (London)* **366,** 690–694.
36. Zarkower, D. & Hodgkin, J. (1992) *Cell* **70,** 237–249.
37. Garriga, G., Guenther, C. & Horvitz, H. R. (1993) *Genes Dev.* **7,** 2097–2109.
38. Harrison, D. A., Gdula, D. A., Coyne, R. S. & Corces, V. G. (1993) *Genes Dev.* **7,** 1966–1978.
39. Schafer, U., Rausch, O., Bouwmeester, T. & Pieler, T. (1994) *Eur. J. Biochem.* **226,** 567–576.
40. Pendleton, J. W., Nagai, B. K., Murtha, M. T. & Ruddle, F. H. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6300–6304.
41. Reuter, G., Giarre, M., Farah, J., Gansz, J., Spierer, A. & Spierer, P. (1990) *Nature (London)* **344,** 219–223.