Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 9287–9291, October 1992
Evolution

Emergence of the keratinocyte growth factor multigene family during the great ape radiation

(fibroblast growth factors/gene amplification/primate evolution)

MICHAEL J. KELLEY*†, MICHAEL PECH‡, HECTOR N. SEUANEZ§¶, JEFFREY S. RUBIN*, STEPHEN J. O'BRIEN§, AND STUART A. AARONSON*||

*Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, MD 20892; ‡F. Hoffmann–La Roche, Ltd., Pharma Division, Preclinical Research, CH-4002 Basel, Switzerland; and §Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, MD 21701-1013

Communicated by Russell F. Doolittle, May 14, 1992

ABSTRACT **The structural gene for human keratinocyte growth factor (KGF), a member of the fibroblast growth factor family, consists of three coding exons and two introns typical of other fibroblast growth factor loci. A portion of the KGF gene, located on chromosome 15, is amplified to ≈16 copies in the human genome, and these highly related copies (which consist of exon 2, exon 3, the intron between them, and a 3′ noncoding segment of the KGF transcript) are dispersed to multiple human chromosomes. The KGF-like sequences are transcriptionally active, differentially regulated in various tissues, and composed of three distinct classes of coding sequences that are 5% divergent from each other and from the authentic KGF sequence. Multiple copies of KGF-like genes were also discovered in the genomic DNAs of chimpanzee and gorilla but were not found in lesser apes (gibbon), Old World monkeys (African green monkey and macaques), mice, or chickens. The pattern of evolutionary occurrence suggests that a primordial KGF gene was amplified and chromosomally dispersed subsequent to the divergence of orangutan from African apes but before the trichotomous divergence of human, chimpanzee, and gorilla 5–8 million years ago. The appearance of a transcriptionally active and chromosomally dispersed multigene KGF family may have implications in the evolution of the great apes and humans.**

The fibroblast growth factors (FGFs) are mitogens that have been implicated in a variety of processes including embryogenesis, angiogenesis, and wound healing (1, 2). Seven distinct members of this growth factor family, in which protein products differ by as much as 67% (3), have been identified in mammalian species (2). This diversity and the presence of multiple gene homologues of FGF in the genomes of species as divergent as *Xenopus* and humans suggest that gene duplications leading to the evolution of the present-day FGF family occurred early in the radiation of vertebrates.

Keratinocyte growth factor (KGF) is the most recently identified member of the FGF family (4)—hence, the alternative designation FGF-7 (5). Unlike other FGFs, which exhibit a broad range of targets, KGF action appears tightly restricted to epithelial cells. The normal expression of KGF by stromal cells of epithelial tissues suggests its physiological role as a major paracrine effector of epithelial cell proliferation (3). Our present studies were undertaken in an effort to characterize the genomic structure of the KGF coding sequence and its chromosomal location, as well as to search for additional FGF family members closely related to KGF. We were able to identify a multigene family of dispersed KGF-like sequences that was generated during the recent evolution of higher primates.

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.

MATERIALS AND METHODS

Nucleic Acid Isolation and Oligonucleotides. Genomic DNAs (6) and total cellular RNA (7) were isolated from cell lines (described in ref. 8) or tissues. Chimpanzee, gorilla, orangutan, and gibbon DNAs were from R. Callahan (National Institutes of Health, Bethesda, MD). Poly(A)$^+$ RNA was isolated from total cellular RNA by using oligo(dT)-cellulose spin columns (Pharmacia) per manufacturer's recommendations. Poly(A)$^+$ RNA from human fetal tissue was purchased from Clontech. Oligonucleotides, synthesized on a Biosearch 8700 DNA synthesizer, were (*i*) PCR primers (Fig. 1*A*): a, ATGCACAAATGGATACTGAC; b, TGTAATTATTCTTCATCTCT; c, ATATCATGGAAATCAGGACA; d, TCTTTGCATAGAGTTTTCCT; e, AAGAATGCAATGAAGATTGT; f, AGTTATTGCCATAGGAAGAA; and (*ii*) probes [see Fig. 2*A*; numbers correspond to nucleotide in the human KGF cDNA sequence (3); oligonucleotides in the reverse direction were composed of antisense sequence corresponding to the indicated nucleotides]: A, 446–494 and 530–481; B, 574–587 and 630–573; C, 631–687 and 731–674; D, 732–790 and 836–777; E, 837–890 and 930–877; F, 931–985 and 1027–972; G, 3748–3807 and 3853–3794.

PCR. PCR was done with 50 mM KCl/10 mM Tris·HCl, pH 8.3 at room temperature/1.5 mM $MgCl_2$/0.01% gelatin/200 μM in each dNTP/1 μM oligonucleotide primers/2.5 units of *Taq* polymerase (Perkin–Elmer) in a 100-μl vol for 25–30 cycles of 30 sec at 94°C, 1 min at 55°C, and 2–4 min at 72°C. Template DNA was 1 μg of genomic DNA, 1–10 ng of KGF cDNA clone 32 (3), or ≈4% of the first-strand cDNA made from 1 μg of poly(A)$^+$ RNA, depending on the application. First-strand cDNA synthesis was done on 1–5 μg of poly(A)$^+$ RNA with a specific KGF primer from the 3′ untranslated region (AGAAGAAAGAAAACAGTCCA) and a cDNA synthesis kit (BRL).

DNA Cloning and Sequencing. PCR products from reactions done with genomic or first-strand cDNA as template and oligonucleotide primers modified by addition of a restriction enzyme site to the 5′ end were digested with the corresponding restriction enzyme(s), electrophoresed through agarose, isolated by GeneClean (Bio 101, La Jolla, CA), and ligated into a pGEMz (Promega) vector. KGF exon 1- and exon 2–3-specific subclones were obtained by cloning the 525-bp *Eco*RI–*Kpn* I and 437-bp *Eco*RI–*Acc* I fragments, respectively, of KGF cDNA clone 32 into a pGEMz vector.

Abbreviations: KGF, keratinocyte growth factor; FGF, fibroblast growth factor.

†Present address: National Cancer Institute–Navy Medical Oncology Branch, Bethesda, MD 20889-5015.
¶Present address: Genetics Section, Instituto Nacional do Cancer, (RJ)/Department of Genetics, Universadade Federal do Rio de Janeiro, Brazil.
||To whom reprint requests should be addressed.

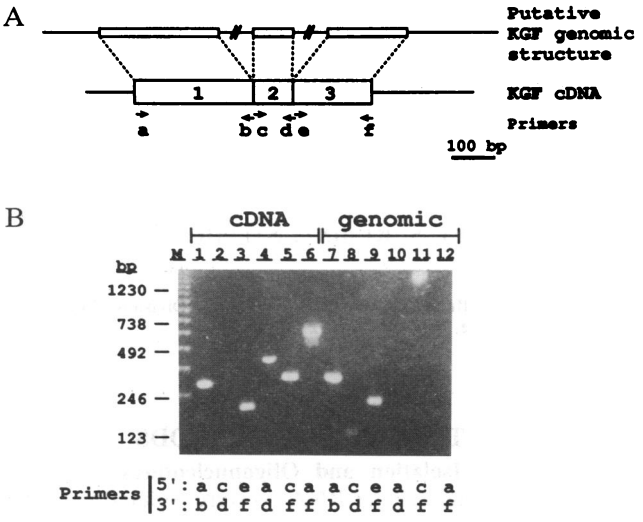

FIG. 1. Determination of human KGF genomic structure by PCR analysis. (*A*) Genomic structure of human KGF gene, as postulated by alignment with *int-2* gene. Numbered open boxes represent coding regions from respective KGF exons. PCR primers are shown as arrows to designate orientation and location relative to KGF cDNA. (*B*) PCRs with KGF cDNA (lanes 1–6) or human genomic DNA (lanes 7–12) as template were primed with different combinations of oligonucleotides (*A*), as indicated below individual lanes, and electrophoresed on an agarose gel containing ethidium bromide. The primer combination c and d produce faintly visible products of ≈100 base pairs (bp) with either cDNA or genomic DNA as template. PCR conditions were as described. Lane M contains a 123-bp DNA ladder as size marker.

DNA sequences were determined by the dideoxynucleotide chain-termination method with Sequenase (United States Biochemical) and were analyzed with the IntelliGenetics Suite programs (IntelliGenetics).

Nucleic Acid Hybridization. Southern blots were done as described (9), except that neutralization of the gel and blotting were done with 1 M ammonium acetate, pH 7.0. Hybridizations were done overnight at 42°C in 5× standard saline/citrate (SSC) (1× SSC is 150 mM NaCl/15 mM sodium citrate)/2.5× Denhardt's solution/10% (wt/vol) dextran sulfate/50% (vol/vol) formamide. Washes were at 50°C in 0.1× SSC. RNA blots were done as described (10). cDNA probes were labeled with [α^{32}P]dCTP with Amersham's nick-translation kit. Oligonucleotide probes were generated by primer-extension reactions and two 60-mers that were annealed via 14 complementary nucleotides at their 3′ ends, according to the method of Studencki and Wallace (11).

RESULTS

Multiple, Highly Conserved Copies of KGF Exons 2 and 3 in the Human Genome. To investigate KGF genomic intron/exon structure, oligonucleotides homologous to the ends of the coding sequence of each predicted exon were synthesized (Fig. 1*A*). The intron/exon junctions were identified by alignment with the published genomic structure of another member of the FGF family, *int-2* (12). When the PCR was done by using pairs of these oligonucleotides with the KGF cDNA (3) as template, PCR products of the predicted size were obtained with all combinations of primers (Fig. 1*B*, lanes 1–6). When pairs of oligonucleotide primers derived from the ends of each predicted coding exon were used with human genomic DNA as template, we observed PCR products corresponding in size to those obtained with the cDNA as template (Fig. 1*B*, lanes 7–9). Thus, the KGF gene was unlikely to have more than three coding exons. In contrast, when oligonucleotide primers from different predicted exons were used in the PCR on the same genomic DNA, either no products were seen (Fig. 1*B*, lanes 10 and 12) or the products were larger than predicted (Fig. 1*B*, lane 11). These results are consistent with the KGF-encoding gene having three coding exons with introns positioned after nucleotides 731 and 835 in the published cDNA sequence (3).

To further define the KGF genomic structure, we digested human genomic DNA with four restriction enzymes that do not cleave within the KGF coding sequence and performed Southern analysis with a series of nonoverlapping, contiguous 100-bp oligonucleotide probes from the coding region (probes A–F), as well as a probe G from the untranslated 3′ end of the transcript (Fig. 2). The results with two restriction enzymes, *Sac* I and *Bgl* I, are shown in Fig. 2. Probes A, B, and C from KGF coding exon 1 each detected a single band with each restriction enzyme, consistent with the existence of a single-copy KGF gene (Fig. 2). In contrast, probes D (exon 2), E and F (exon 3), and G (3′-untranslated region) each detected multiband patterns of DNA fragments, demonstrating the existence of at least four genomic copies of sequences highly related to KGF exons 2 and 3. Moreover, the intensity of each band observed with a given probe was not uniform, suggesting that the more intense bands might contain more than one copy of KGF exon 2 and 3 sequences. A similar multiband pattern was detected with oligonucleotide probe E (Fig. 2) in *Sac* I-digested genomic DNAs of around 30 unrelated individuals (data not shown). Thus, multiple KGF-like sequences are present in all human DNAs tested, and the presence of these sequences cannot be explained by polymorphism.

To further characterize these sequences, we PCR-cloned the 1.5-kilobase (kb) genomic fragments with primers derived from KGF sequences at the 5′ and 3′ ends of exon 2 and 3, respectively (primers c and f in Fig. 1*A*) into a pGEMz vector. Nucleotide sequence analysis of the exon 3 region of 23 clones derived from two independent PCRs on the same human genomic DNA revealed that these clones could be arranged into four distinct groups. One representative from each group was then sequenced in the exon 2 and intron regions. The sequence differences in exons 2 and 3, as well as in nucleotides of the intron, are depicted in Fig. 3*A*.

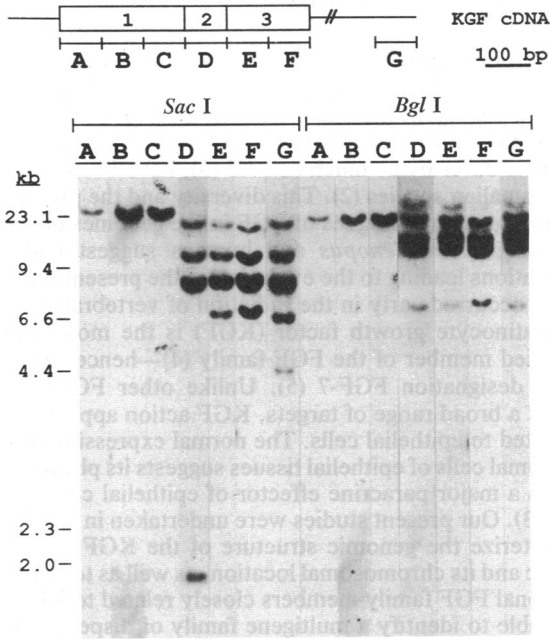

FIG. 2. Southern blot analysis of KGF-like sequences. Human genomic DNA was digested with *Sac* I or *Bgl* I, fractionated on a 1.0% agarose gel, transferred to nitrocellulose, and hybridized with oligonucleotide probes A–G derived from different regions of the KGF cDNA. DNA digests with *Bam*HI and *Eco*RI gave similar hybridization patterns.

Differences in nucleotide sequence between the four prototypes (I–IV) were ≈5%. However, the exon 3 sequences of clones within groups also varied but by only ≈0.2% (Fig. 3*A*). Whether this degree of sequence variation is due to *Taq* polymerase infidelity during PCR or whether groups II–IV contain more than a single member is not resolved.

The sequence of prototype I was identical to exons 2 and 3 of the KGF cDNA, establishing it as a PCR clone of the KGF gene. The KGF open reading frame was maintained in each of the other prototypes as well (Fig. 3*B*). For prototype IV, a 1-bp deletion at nucleotide position 998 at the end of exon 3 would result in replacement of the C-terminal 10-amino acid residues with 7 different amino acids (Fig. 3*B*). Each prototype genomic clone also contained 14 bp just upstream from exon 3 (Fig. 3*A*) that was consistent with a 3′ splice site (13), suggesting that the intron was functional in each set.

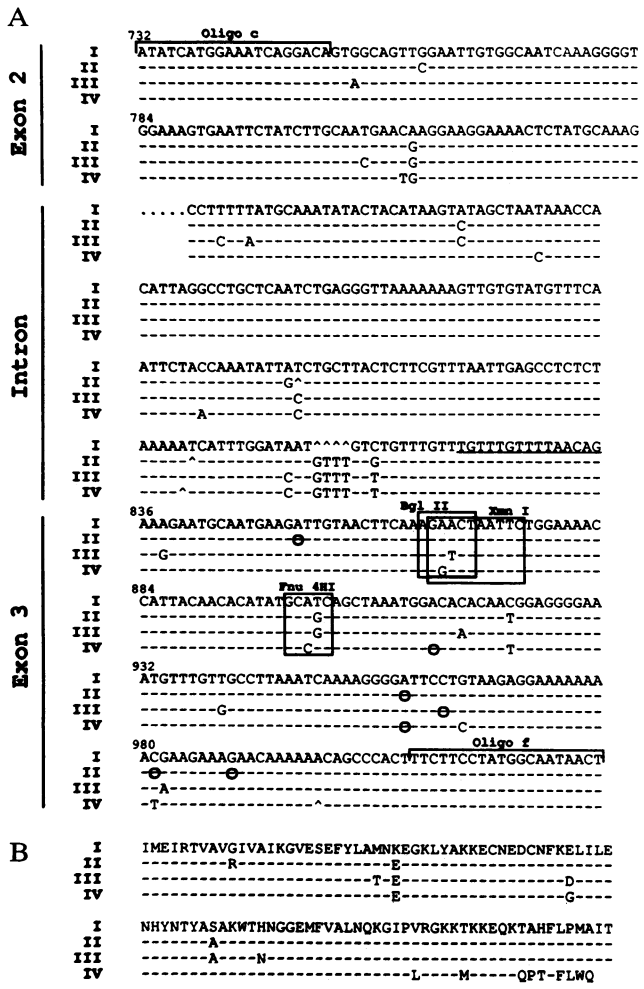

FIG. 3. Comparison of nucleotide and amino acid sequences of human KGF family. (*A*) Consensus sequences of 23 human genomic PCR clones for prototypes II–IV are compared with prototype I sequence, identical to KGF. The number of clones of prototypes I, II, III, and IV that were sequenced in exon 3 were three, eight, six, and six, respectively, of which three, four, four, and five, respectively, were identical. The PCR primers c and f used for amplification of exon 2–3 sequences are located as indicated but were not included for sequence-divergence calculations. Exon 2 and intron sequences (192 bp) immediately adjacent to exon 3 were determined for one clone of each prototype; dots (. . . .) represent ≈1000 bases of the intron not sequenced. Identical nucleotides are depicted by a hyphen (-); gaps in the sequence are denoted by a carat (^). The boxed regions contain areas of restriction-site differences that distinguish the four prototypes. *Bgl* II, *Xmn* I, and *Fnu*4HI have recognition sites in prototypes III, I and II, and II and III, respectively. The splice-acceptor site of the intron is underlined. Circled hyphens in exon 3 represent nucleotides that vary between clones of a prototype. These nucleotides are identical to the corresponding KGF (prototype I) nucleotide in all but one clone of a prototype, except for prototype III nucleotide 963, which was thymine in two clones. Nucleotide numbering corresponds to the KGF cDNA (3). (*B*) Predicted amino acid sequence of prototype I (KGF) exons 2 and 3 is shown. Corresponding sequences of prototypes II–IV are aligned, indicating maintenance of the open reading frame in each. Hyphens represent identity.

Expression of Spliced Prototype II and III KGF-Like mRNAs by Human Cells. The four prototypes of KGF-like sequences can be distinguished by the restriction pattern generated with *Bgl* II, *Fnu*4HI, and *Xmn* I (see Fig. 3*A*). The four predicted transcripts were sought in various tissue and cell-line RNAs by digesting the PCR products generated with oligonucleotides c and f (Fig. 1*A*) as primers and first-strand cDNA as template with these three restriction enzymes. Transcripts distinct from KGF (prototype I) were detected in fetal brain, fetal kidney, primary human umbilical vein endothelial (HUVE) cells, and tumor cell lines A101D (malignant melanoma), A388 (squamous cell carcinoma), A431 (vulvar squamous cell carcinoma), and A498 (renal cell carcinoma). Sequence analysis of PCR-generated cDNAs cloned from fetal brain and kidney directly showed that these cDNAs were derived from prototype II and III transcripts. Absence of the intron between exons 2 and 3 in these cDNAs further indicated that these KGF-like RNAs were spliced correctly (Fig. 4*A*) and could not have resulted from amplification of contaminating genomic DNA. The KGF open reading frame was maintained in these prototype II and III cDNAs.

To determine whether the sizes of prototype II and III transcripts could be distinguished from that of KGF and evaluate their relative levels of expression, we subjected poly(A)$^+$ RNAs from these same tissues and cell lines to RNA blot analysis. As shown in Fig. 4*B*, the human umbilical vein endothelial cells, as well as the tumor cell lines in which prototype II and III transcripts but not the prototype I (KGF) transcript were specifically detected by PCR analysis, showed low levels of 2.0- and 2.4-kb RNA species. The smaller transcript could be readily distinguished from the major 2.4-kb KGF transcript observed in M426 cells. An exon 1 probe detected the KGF transcript in M426 cells with a sensitivity greater than that of the exon 3 probe (Fig. 4*B*). However, the exon 1 probe failed to hybridize to either the 2.4- or 2.0-kb transcripts in cells expressing prototype II and

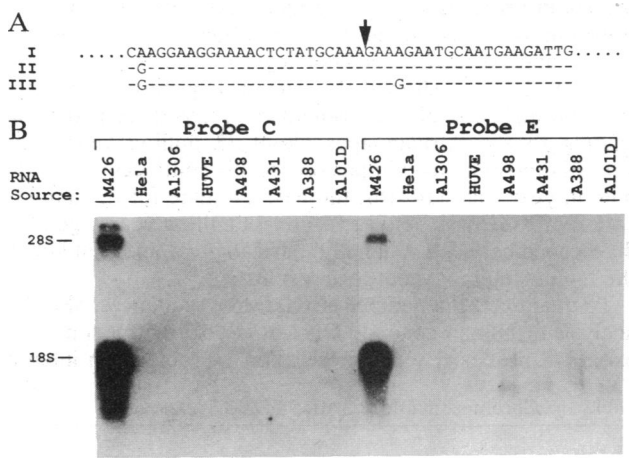

FIG. 4. Detection of KGF-like transcripts in human cells. (*A*) Comparison of nucleotide sequences of prototype II and III cDNA clones obtained by PCR analysis of human fetal brain mRNA with the sequence of KGF cDNA (prototype I). The junction of exons 2 and 3 is shown by an arrow. Hyphens represent identity. (*B*) Around 10 μg of poly(A)$^+$ RNA from various cell types was subjected to RNA blotting as described. Duplicate blots were hybridized with oligonucleotide probe C or E (Fig. 2). Positions of 28S and 18S RNA are indicated. HUVE, human umbilical vein endothelial cells.

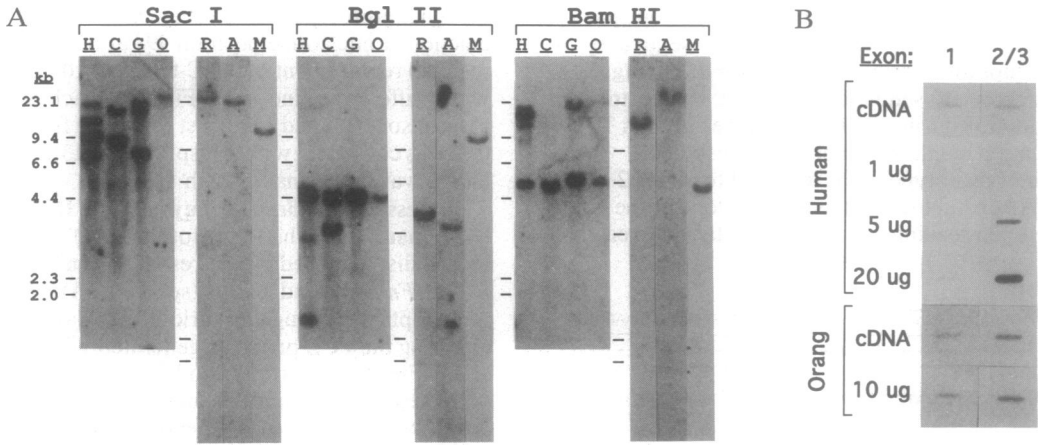

FIG. 5. Quantitation of KGF genomic sequences in great apes and humans. (*A*) Human (H), chimpanzee (C), gorilla (G), orangutan (O), rhesus (R), African green monkey (A), and mouse (M) genomic DNAs (5–20 μg) were digested with *Sac* I, *Bgl* II, and *Bam*HI and subjected to Southern blotting as described. The blot was hybridized with oligonucleotide probe D (Fig. 2). No additional bands were detected with reduced-stringency hybridization. (*B*) Serial dilutions of KGF cDNA and human genomic DNA were subjected to slot-blot analysis on a Minifold II apparatus (Schleicher & Schuell) by using either a KGF exon 1 or exon 2–3 probe. Representative autoradiograms with comparable signal intensity for each probe hybridized to 100 pg of KGF cDNA clone 32 (3) are shown. Densitometry of the signal from 1, 5, or 20 μg of human genomic DNA (using a Bio-Rad model 620 video densitometer) indicated that the signal generated by the exon 2–3 probe was ≈16 times more intense than the signal from the exon 1 probe. Representative autoradiograms from a similar analysis with 10 pg of KGF cDNA clone 32 and 10 μg of orangutan (Orang) genomic DNA are shown. Densitometry of the orangutan genomic DNA indicated that the signal generated by exon 2–3 probe was approximately equal to the signal from exon 1 probe.

III transcripts (Fig. 4*B*). These findings directly established the absence of KGF exon 1 sequences in transcripts of prototypes II and III KGF-like genes.

Amplification of KGF-Encoding Genes During the Great Ape Radiation. The high degree of sequence identity (95% nucleotide matching) between human KGF and the additional KGF-like segments suggested that the amplification event that created them was rather recent. To estimate the timing of the amplification, we used exon 2 (probe D)- and exon 3 (probe E)-specific probes to verify gene multiplication in the closest relatives to humans, the great apes. Both of these probes revealed multiple KGF segments in DNA from chimpanzee (*Pan troglodytes*) and gorilla (*Gorilla gorilla*) with three restriction enzymes that fail to cut within the human coding exons (shown for probe E in Fig. 5*A*). In contrast, a single fragment was revealed with the same enzymes and probes in orangutan (*Pongo pygameus*). Similarly, a single fragment was seen in lesser ape (gibbon; data not shown), rhesus and African green monkeys, mouse (Fig. 5*A*), and chicken (data not shown). Because it is generally accepted that the ancestor of the orangutan diverged from the line leading to the African apes about 13 million years ago, whereas the human–chimpanzee–gorilla split occurred ≈5–8 million years ago (14–16), the best estimate for the amplification of KGF was between 8 and 13 million years ago after divergence between Asian and African great apes but before the human–chimpanzee–gorilla split.

To quantitate the number of KGF-like sequences, slot-blot analysis of human genomic DNA was done with exon 1 and exon 2–3 probes. By using the exon 1 probe as an internal control, exon 2, exon 3, and the 3′ untranslated region were estimated to be present at ≈16 copies per human haploid genome (Fig. 5*B*). Similar analysis was done with chimpanzee, gorilla, orangutan, and gibbon. More than one copy of KGF-like sequences were also found only in chimpanzee and gorilla (data not shown) but were not found in orangutan (Fig. 5*B*), supporting the concept that the event(s) that led to the generation of the KGF family in humans occurred during speciation of the great apes.

Dispersed KGF-Like Sequences in the Human Genome. To identify the chromosomal location of the KGF gene, genomic DNAs of rodent × human somatic-cell hybrids (17, 18) were digested with *Sac* I and analyzed for the presence of the diagnostic 25-kb restriction fragment revealed by an exon 1 probe (see Fig. 2). In a panel of 45 hybrid cell lines, this fragment was 95% concordant with chromosome 15 and 18–65% discordant with the other human chromosomes, indicating that KGF is located on chromosome 15 (Table 1). An exon 3 probe (oligonucleotide E in Fig. 2) recognizes four human *Sac* I fragments of size 23 kb, 15 kb, 9 kb, and 7 kb (Fig. 2, *Sac* I digestion, probe E). Analysis of the same series of somatic cell-hybrid DNAs with this probe localized the 23-kb fragment to chromosome 15, but it was not possible to assign the 15-, 9-, or 7-kb bands that appeared to be present at >1 copy number (Fig. 2) to any one chromosome. For every human chromosome the discordance with these three restriction fragments was >25%. As this panel has been diagnostic for gene mapping of >100 human loci on each of the 23 chromosomes (19), we interpret the high discordance to reflect multiple KGF-like sequences dispersed to several human chromosomes.

Table 1. Chromosomal localization of KGF

	Chromosome																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
Concordance, no.	26	27	30	25	29	22	27	26	26	25	25	30	28	25	39	26	27	28	23	31	18	27	15
Discordance, no.	8	6	9	12	8	16	10	11	10	11	16	8	10	14	2	11	10	10	13	9	18	12	27
Total, no.	34	33	39	37	37	38	37	37	36	36	41	38	38	39	41	37	37	38	39	40	36	39	42
Discordance, %	24	18	23	32	22	43	27	30	28	31	39	21	26	36	5	30	27	26	33	23	50	31	64

DNAs of a panel of 45 genetically characterized somatic cell hybrids, as well as DNAs from rodent and human parental cell lines, were digested with *Sac* I and subjected to Southern-blot analysis using a KGF oligonucleotide C (exon 1) probe. The human 25-kb band was readily distinguishable from the mouse and hamster bands. The results of this analysis localize KGF to chromosome 15.

DISCUSSION

The FGF family is represented by at least seven divergent genes, from which each characterized product acts as a potent mitogen (1, 2). Homologues of several members of this family exist in species as divergent as *Xenopus* (ref. 20 and M.J.K., unpublished observation) and human (1, 2), indicating that amplification events leading to the present-day members of this family have preceded the divergence of mammals and reptiles. Our present studies demonstrate that the KGF gene, a recently discovered member of this family, is organized into three coding exons, analogous in structure to those of other characterized FGF members (12, 21–25). In addition, a portion of the KGF gene encompassing exons 2 and 3, as well as the corresponding intron and 3′ untranslated region was amplified in human, chimpanzee, and gorilla. Evidence that these events occurred recently in the evolution of primates was provided by the detection of only a single KGF gene in orangutan, gibbon, African green monkey, rhesus monkey, and all the nonprimate species analyzed. Amplification of the KGF gene likely occurred after the divergence of orangutans from the human/chimpanzee/gorilla clad ≈13 to 16 million years ago and before the divergence of human, chimpanzee, and gorilla from their common ancestor some 5 to 8 million years ago (14–16).

By means of somatic cell-hybrid analysis, it was possible to map the gene for KGF to chromosome 15. In contrast, basic FGF and FGF-5 are located at 4q25 (26) and 4q21 (27), respectively, whereas *hst* and *int-2* are clustered within 45 kb at 11q13 (27, 28), consistent with independent amplifications by a mechanism involving unequal crossover on these respective chromosomes. Because the gene-amplification events involving KGF occurred during a span of primate evolution in which full chromosomal duplications have not occurred (29, 30), the dispersion of KGF-like sequences cannot reflect this mechanism.

The majority of described pseudogenes are intronless and, thus, appear to have been generated through reverse transcription of a spliced RNA intermediate (19, 31, 32). In contrast, the KGF-like sequences all contain introns. The generation of a large number of dispersed intron-containing pseudogenes from a single functional gene is unusual, and we know of no example of any characterized growth factor for which pseudogenes have been reported. Thus, if the KGF-like sequences are pseudogenes, they would be distinctive in several respects.

Rapid gene amplification in cultured mammalian cells, tumors, or during development of invertebrates has been frequently documented and is thought to occur by a variety of mechanisms including segregation and replication-driven processes (for review, see ref. 33). Overexpression of amplified genes, such as those responsible for drug resistance, confers a selective advantage. Analogous mechanisms have been implied in the generation of multigene families, such as the histone (34) and immunoglobulin genes (35). In the well-documented example of the human immunoglobulin variable region κ genes, expansion in family size contributes to the diversity of the antibody repertoire. During κ locus evolution, variable-region κ-encoding genes may have behaved as transposons (36) with some members being dispersed to other chromosomes (37). These orphons also display a high degree of sequence identity [97% (38)]. Whether the amplification of KGF was driven by selective processes or simply represents a frozen accident of stochasticity in the evolution of the mammalian genome remains to be seen.

1. Burgess, W. H. & Maciag, T. (1989) *Annu. Rev. Biochem.* **58**, 575–606.
2. Benharroch, D. & Birnbaum, D. (1990) *Isr. J. Med. Sci.* **26**, 212–219.
3. Finch, P., Rubin, J., Miki, T., Ron, D. & Aaronson, S. (1989) *Science* **245**, 752–755.
4. Rubin, J. S., Osada, H., Finch, P. W., Taylor, W. G., Rudikoff, S. & Aaronson, S. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 802–806.
5. Baird, A. & Klagsbrun, M. (1991) *Ann. N.Y. Acad. Sci.* **638**, xiii–xvi.
6. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3**, 2303–2308.
7. Glisin, V., Crkvenjakov, R. & Byus, C. (1974) *Biochemistry* **13**, 2633–2637.
8. Giard, D. J., Aaronson, S. A., Todaro, G. J., Arnstein, P., Kersey, J. H., Dosik, H. & Parks, W. P. (1973) *J. Natl. Cancer Inst.* **51**, 1417–1423.
9. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
10. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 7.39–7.52.
11. Studencki, A. B. & Wallace, R. B. (1984) *DNA* **3**, 7–15.
12. Brookes, S., Smith, R., Casey, G., Dickson, C. & Peters, G. (1989) *Oncogene* **4**, 429–436.
13. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
14. Andrews, P. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51**, 419–428.
15. Janczewski, D. N., Goldman, D. & O'Brien, S. J. (1990) *J. Hered.* **81**, 375–387.
16. Sibley, C. G. & Ahlquist, J. E. (1987) *J. Mol. Evol.* **26**, 99–121.
17. O'Brien, S. J., Nash, W. G., Goodwin, J. L., Lowy, D. R. & Chang, E. H. (1983) *Nature (London)* **302**, 839–842.
18. O'Brien, S. J., Bonner, T. I., Cohen, M., O'Connell, C. & Nash, W. G. (1983) *Nature (London)* **303**, 74–77.
19. O'Brien, S. J. & Johnson, P. A. (1990) in *Genetic Maps*, ed. O'Brien, S. J. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 5.197–5.210.
20. Kimelman, D. & Kirschner, M. (1987) *Cell* **51**, 869–877.
21. Abraham, J. A., Whang, J. L., Tumolo, A., Mergia, A., Friedman, J., Gospodarowicz, D. & Fiddes, J. C. (1986) *EMBO J.* **5**, 2523–2528.
22. Yoshida, T., Miyagawa, K., Odagiri, H., Sakamoto, H., Little, P. F., Terada, M. & Sugimura, T. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7305–7309.
23. Zhan, X., Bates, B., Hu, X. G. & Goldfarb, M. (1988) *Mol. Cell. Biol.* **8**, 3487–3495.
24. Mergia, A., Tischer, E., Graves, D., Tumolo, A., Miller, J., Gospodarowicz, D., Abraham, J. A., Shipley, G. D. & Fiddes, J. C. (1989) *Biochem. Biophys. Res. Commun.* **164**, 1121–1129.
25. Marics, I., Adelaide, J., Raybaud, F., Mattei, M. G., Coulier, F., Planche, J., de Lapeyriere, O. & Birnbaum, D. (1989) *Oncogene* **4**, 335–340.
26. Howard Hughes Medical Institute Human Gene Mapping (1990) *Cytogenet. Cell Genet.* **55**, 1–786.
27. Nguyen, C., Roux, D., Mattei, M., de Lapeyriere, O., Goldfarb, M., Birnbaum, D. & Jordan, B. R. (1988) *Oncogene* **3**, 703–708.
28. Adelaide, J., Mattei, M. G., Marics, I., Raybaud, F., Planche, J., de Lapeyriere, O. & Birnbaum, D. (1988) *Oncogene* **2**, 413–416.
29. Yunis, J. J. & Prakash, O. (1982) *Science* **215**, 1525–1530.
30. Seuanez, H. N. (1987) in *Cytogenetics: Basic and Applied Aspects*, eds. Obe, G. & Basler, A. (Springer, Berlin), pp. 65–89.
31. Anagnou, N. P., Antonarkis, S. E., O'Brien, S. J., Modi, W. S. & Nienhuis, A. W. (1988) *Am. J. Hum. Genet.* **42**, 491–509.
32. Anagnou, N. P., O'Brien, S. J., Shimada, T., Nash, W. G., Chen, M. & Nienhius, A. W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5170–5174.
33. Stark, G. R., Debatisse, M., Giulotto, E. & Wahl, G. M. (1989) *Cell* **57**, 901–908.
34. Childs, G., Maxson, R., Cohn, R. H. & Kedes, L. (1981) *Cell* **23**, 651–663.
35. Honjo, T. & Habu, S. (1985) *Annu. Rev. Biochem.* **54**, 803–830.
36. Pech, M., Jaenichen, H. R., Pohlenz, H. D., Neumaier, P. S., Klobeck, H. G. & Zachau, H. G. (1984) *J. Mol. Biol.* **176**, 189–204.
37. Borden, P., Jaenichen, R. & Zachau, H. G. (1990) *Nucleic Acids Res.* **18**, 2101–2107.
38. Loetscher, E., Zimmer, F.-J., Klopstock, T., Grzeschik, K.-H., Jaenichen, R., Straubinger, B. & Zachau, H. G. (1988) *Gene* **69**, 215–223.