*Proc. Natl. Acad. Sci. USA*
Vol. 88, pp. 671–675, January 1991
Biochemistry

# Identification and characterization of "zinc-finger" domains by the polymerase chain reaction

(DNA amplification/gene family/metal-binding domain)

GERALD R. PELLEGRINO* AND JEREMY M. BERG†‡

*Department of Biology and †Department of Chemistry, The Johns Hopkins University, Baltimore, MD 21218; and ‡Department of Biophysics and Biophysical Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205

*Communicated by Daniel Nathans, September 21, 1990 (received for review August 13, 1990)*

ABSTRACT **We have developed a method for amplifying DNA fragments containing tandem arrays of "zinc-finger" sequences of the transcription factor IIIA ($Cys_2$-$His_2$) type by using the polymerase chain reaction. Because these sequences occur as tandem arrays, a ladder of bands is produced upon amplification using primers derived from the amino- and carboxyl-terminal sequences of a zinc-finger domain. The "rungs" of this ladder correspond to DNA fragments encoding one zinc-finger domain, two adjacent zinc-finger domains, and so on. This is demonstrated by isolating individual bands corresponding to *n* zinc-finger domains and reamplifying them with the same primers. This yields a band of the original size as well as bands corresponding to 1 through ($n - 1$) zinc-finger domains. Direct evidence that these bands encode zinc-finger domains was obtained by cloning and sequencing a collection of the amplification products. Due to the lack of redundancy in the sequences obtained, we conclude that each band corresponds to a large number of unique zinc-finger-encoding sequences. The results from amplification reaction mixtures using genomic DNA from a variety of sources as template provide further evidence that zinc-finger domains occur widely and frequently in eukaryotic genomes. We believe that this method is a powerful technique for the isolation and characterization of zinc-finger-encoding genes.**

The development of methods that allow the isolation of genes encoding homologs of characterized proteins has greatly facilitated the elucidation of the composition of gene families and the identification of candidates for which functions can be rationally hypothesized based on the presence of characteristic sequences. Among the more powerful methods for isolation of such genes are those based on low-stringency nucleic acid hybridization. However, these methods suffer from several drawbacks including limitations on probe degeneracy and the labor-intensive processes required for clone verification. Methods based on the polymerase chain reaction (PCR) have been developed to facilitate detection of homologous sequences (e.g., see ref. 1). The major requirement of the PCR-based methods is the identification of two short stretches of sequence that are relatively well conserved within a gene family.

A number of gene families have been discovered that encode proteins that contain domains that are internally repeated, often tandemly. For example, many proteins involved in blood coagulation have one or more copies of the so-called Kringle sequence, an ≈80-amino acid sequence containing three conserved disulfide bonds. This motif occurs once in urokinase and factor XII, twice in prothrombin and tissue plasminogen activator, 5 times in plasminogen, and a total of 38 times in a cDNA clone of apolipoprotein (a) (2). Our interests lie in studying the structures and functions of another class of proteins that contain tandemly repeated domains, namely, the zinc-finger proteins of the $Cys_2$-$His_2$ type (3–5). These proteins are characterized by the presence of one or more sequences that approximate the consensus (Tyr/Phe)-Xaa-Cys-$Xaa_m$-Cys-$Xaa_3$-Phe-$Xaa_5$-Leu-$Xaa_2$-His-$Xaa_o$-His-$Xaa_n$, where $m$ is 2 or 4, $n$ is often 5, and $o$ is 3 or 4. Each of these sequences appears to form a structural domain organized around a zinc ion coordinated to the invariant cysteine and histidine residues. The presence of tandem sequences of this form was discovered in a cDNA clone of *Xenopus* transcription factor IIIA (TFIIIA) (6, 7). Subsequently, a large number of other members of this family have been discovered. Many additional members have been discovered independently, that is, by examination of a sequence that was identified for other reasons without the intention of looking for zinc-finger domains. Other members have been identified intentionally by using low-stringency hybridization methods. These include large groups identified in mouse (8), *Xenopus* (9), and human cDNA libraries (10, 11). We now report PCR-based methods that appear to be quite powerful in identifying zinc-finger-encoding sequences. The presence of tandem arrays of similar sequences results in some unusual behavior that can be used to advantage in identifying appropriate fragments. Our results represent further evidence of the widely occurring nature of TFIIIA-like zinc-finger sequences in eukaryotic genomes (12, 13) and provide a method for isolating putative zinc-finger-encoding sequences for further characterization.

## MATERIALS AND METHODS

**Amplification Primers.** The PCR primers were synthesized on the Cyclone DNA synthesizer (Milligen, Bedford, MA/Biosearch) and purified on oligo purification cartridges (Applied Biosystems) according to the manufacturer's instructions. Primers derived from the carboxyl terminus of the TFIIIA-like zinc-finger domain were designed to incorporate an *Eco*RI site into the amplification product and primers from the amino terminus incorporated a *Bam*HI site into the product to facilitate cloning. The amino acid sequence on which the primer sequences were based is derived from a consensus of 131 zinc-finger sequences (B. A. Krizek, B. T. Amann, V. J. Kilfoil, D. L. Merkle, and J.M.B., unpublished observations) (for the sequence, position, and degeneracy of each primer, see Table 1).

**PCR Conditions.** The amplification reactions were carried out in the presence of *Thermus aquaticus* polymerase (Perkin–Elmer/Cetus) with the buffer, deoxynucleoside triphosphates, and primer concentrations suggested by the manufacturer. From 2 μg to 8 μg of genomic template was used per reaction mixture. Each reaction was allowed to cycle 30 times with each cycle as follows: 1 min at 94°C, 2 min at 45°C,

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: PCR, polymerase chain reaction; TFIIIA, transcription factor IIIA.

Table 1. Amplification primers

| Primer | Sequence | Degeneracy |
|---|---|---|
| Cys-Xaa$_2$-Cys | 5′-GTA-GAA-TTC-TG(T,C)-CCN-GA(G,A)-TG(T,C)-GGN-AA(G,A)-TCN-TT-3′ *Eco*RI | $2^{10}$ |
| Cys-Xaa$_4$-Cys | 5′-GTA-GAA-TTC-TG(T,C)-CCN-NTN-GA(GA)-GGN-TG-3′ *Eco*RI | $2^{10}$ |
| H-C Link | 5′-GTC-GGA-TCC-ANG-G(T,C)T-T(T,C)T-CNC-CNG-T(A,G)T-G-3′ *Bam*HI | $2^9$ |
| His-Xaa$_3$-His | 5′-GTC-GGA-TCC-(A,G)TG-N(A,G)(T,C)-NC(T,G)-(C,T)TG-(A,G)TG-3′ *Bam*HI | $2^{10}$ |

Sequence of each primer used is shown along with the name by which it is referred to in the text. Positions at which a degeneracy was introduced into the oligonucleotide sequence are shown in parentheses (N signifies that all four nucleotides were used). At degenerate positions, an equimolar ratio of the indicated nucleotides was used in the synthesis. Restriction sites, added to facilitate cloning of the amplification products, are underlined. The degeneracy of each primer is indicated.

and 3 min at 72°C. Each reaction was then kept at 72°C for 10 min. The products were separated on agarose gels and visualized with ethidium bromide. For further analysis, individual bands were excised from the agarose gel and the amplification product was electroeluted into 10 M ammonium acetate and precipitated with ethanol. These fragments were then used as the template and amplified as described above.

**Cloning and Sequencing of Reaction Products.** Amplification products were gel-purified and electroeluted as described above. Fragments were then digested with *Eco*RI and *Bam*HI overnight and gel-purified on 3% (wt/vol) NuSieve (FMC) agarose gels. Individual bands were excised and ligated into the pEMBL vector (14) overnight at room temperature. These constructs were transformed into DH5$\alpha$ competent cells and selected using $\beta$-galactosidase activity as a marker. The recombinant plasmids were then isolated using a modified STET preparation (15) and sequenced as a double-stranded template, in both directions, by using Sequenase (United States Biochemical).

## RESULTS AND DISCUSSION

The PCR depends on the presence of two primer binding sites within a relatively short distance of one another (16). The use of primers derived from sequences at the amino and carboxyl termini of a sequence motif that occurs in tandem arrays should lead to amplification of a family of sequences as illustrated in Fig. 1. The primer derived from the carboxyl terminus may anneal to the appropriate region within the same motif as the amino-terminus-derived primer or in another sequence within the array. This analysis leads to the prediction that a ladder of bands should be produced upon analysis by gel electrophoresis that corresponds to one motif, two adjacent motifs, and so on. We designed primers based on a consensus sequence derived from a data base of 131 zinc-finger sequences (B. A. Krizek, B. T. Amann, V. J. Kilfoil, D. L. Merkle, and J. M. Berg, unpublished observations). For the amino-terminal primer, the consensus amino acid sequence used was Cys-Pro-Glu-Cys-Gly-Lys-Ser-Phe. A $2^{10}$-fold degenerate primer was synthesized using this sequence and allowing for the full degeneracy of the genetic code. Similarly, carboxyl-terminal primers were designed based on the consensus sequence His-Gln-Arg-Thr-His and on the sequence of the H-C link His-Thr-Gly-Glu-Lys-Pro, which often connects adjacent zinc-finger sequences (13). These primers will be referred to as the Cys-Xaa$_2$-Cys primer, the His-Xaa$_3$-His primer, and the H-C link primer, respectively. The result of amplifying human genomic DNA with appropriate pairs of these primers is shown in Fig. 2. A ladder of bands was generated with the sizes expected for tandemly repeated zinc-finger domains. Furthermore, the Cys-Xaa$_2$-Cys to H-C link fragments are larger than the Cys-Xaa$_2$-Cys to His-Xaa$_3$-His fragments by 15 base pairs (bp) as expected.

Two approaches were taken to confirm that the amplified bands are, in fact, zinc-finger-encoding fragments. (*i*) If the larger molecular weight bands do contain tandem zinc-finger-encoding sequences, then reamplification of a single band by PCR would be expected to yield a ladder that contains the original-sized fragment as well as fragments containing smaller numbers of zinc-finger-encoding units but not larger ones. As shown in Fig. 3, this expectation is borne out by experiment. Importantly, the fact that larger products were not observed eliminates the possibility that the original ladder of bands was an artifact generated by priming one fragment by another. Such priming should be minimal since the fragments generated have non-zinc-finger-related sequences on their 3′ ends derived from the original primers used. This "two-dimensional PCR" provides a simple method to determine if an amplified fragment contains the expected internally repeated sequences.

The second method used to confirm the nature of the amplified sequences involved direct cloning and sequencing of the amplification products. This also allowed us to answer another important question: Does each band represent a

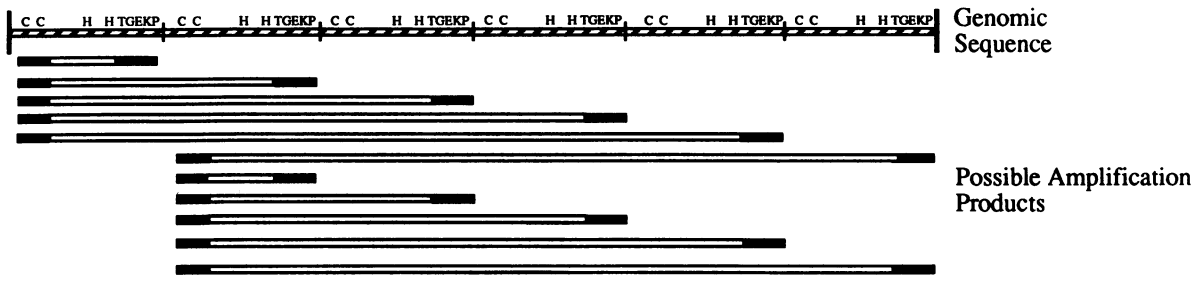


FIG. 1. Amplification of tandemly repeated sequences. The predicted products obtained when a genomic sequence encoding tandemly repeated domains is amplified using the PCR are shown. In this example, the genomic sequence (hatched bar) encodes six zinc-finger domains. Although only a limited number of possible amplification products are shown in this schematic, it is apparent that a molecular size "ladder" of fragments would be obtained from this type of reaction. Amplification primers are represented by a solid bar and the amplified sequences are presented by an open bar. Conserved amino acids in the finger motif are shown above the genomic sequence.

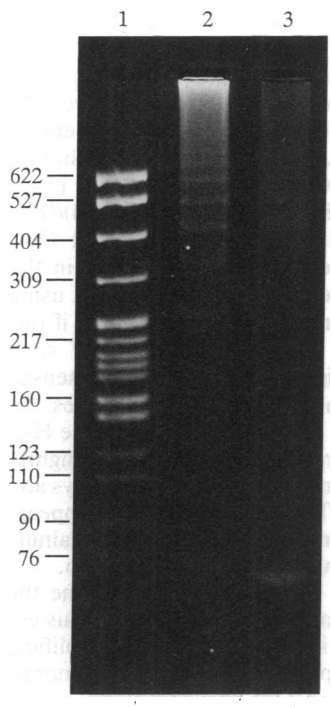


FIG. 2. Amplification of putative zinc-finger-encoding domains from the human genome. Upon the amplification of sequences in the human genome with either the Cys-Xaa$_2$-Cys to the His-Xaa$_3$-His primers (lane 2) or the Cys-Xaa$_2$-Cys to H-C link primers (lane 3), a molecular weight ladder of products is obtained. The predicted size of a fragment that encodes one-finger domain amplified using the Cys-Xaa$_2$-Cys and the His-Xaa$_3$-His primers is 81 bp. The fragment in the Cys-Xaa$_2$-Cys to H-C link lane that runs at ≈67 bp is believed to be a primer-derived artifact. Lane 1 contains pBR322 DNA digested with *Msp* I. Positions of size markers in base pairs are shown to the left.

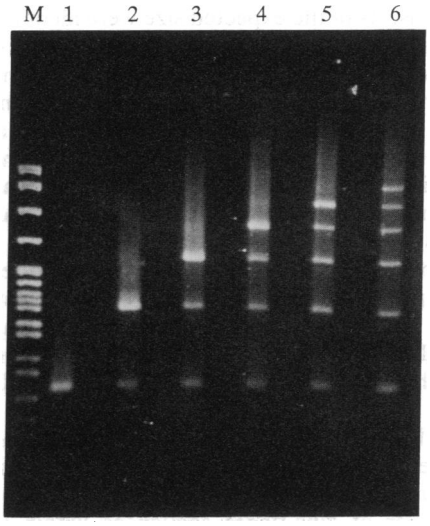


FIG. 3. Two-dimensional PCR. The presence of tandemly repeated sequences in the amplification products was verified by the reamplification of a given sized fragment, as described in the text. The lane labels correspond to the number of zinc-finger-encoding domains contained in the fragment prior to reamplification. Each fragment produces itself and all of the possible smaller size fragments. The Cys-Xaa$_2$-Cys and H-C link primers were used for these reactions. Lane M contains the same molecular size markers described in Fig. 2.

single sequence or a large number of sequences of the same length? The primers used were designed to include restriction sites to allow direct cloning of the products. The band from the Cys-Xaa$_2$-Cys to H-C link primed reaction corresponding in size to that expected for a four-zinc-finger-encoding fragment was isolated and cloned. Colonies that contained plas-

```
C PE C GKSFSKSYSLND H WRL H TGEKP YE C RE C WKSFRQSSSLIQ H RRV H TAVRP
                                                                    Z4FPCR.H1
HE C DE C GKLFSNLSNLIK H RRV H TGERP YE C SE C GKSFSQRSALLQ H RGV H TGEKP

C PE C GKSFLRKSDLTQ H QRV H TGEKP YV C QL Y VSIFRQVSVLIQ H QRV H TGERP
                                                                    Z4FPCR.H2
YE C SE C GKSFSHSTNLYR H RSA H TSTRP YE C SE C GKSFSHSTNLFR H WRV H TGEKP

C PE C GKSFSQKSSLQC H HIL H TGEKP YL C EE C DNVYIRRSHLGR H RKI H TGEGS
                                                                    Z4FPCR.H3
YL C LV C DLVFRSDSYLAE H QRV H TGEKP YL C NL C GWSFSRKSSLEY H HTL H TGEKP

C PE C GKSFLYSSTFSS H KRS H TGEKP YK C EE C GKALLHPQPLGK H EII H TGKKP
                                                                    Z4FPCR.H4
YK C EE C GKAFNQSSSLTK H KKI H TGEKP YK C EE C GKAFNQSSSLTK H KKI H TGEKP

C PE C GKSFINKSNLRL H QRT H TGEKP YE C NE C GKTFHRKSFLTI H QWT H TGEKP
                                                                    Z4FPCR.H5
YE C NE C GKTFRCKSFLTV H QRT H AREKP YA C NE C GKTYSHKSYLTV H HRT H TGEKP

C PE C GKSFSINEKLIW H QRL H SEGKP FK C VE C GKSFSYSSHYIT H QHI H SGEKP
                                                                    Z4FPCR.H6
I S VR C VKAFSVNGSLSR H QRI H TGEKP YQ C KE C GNGFSCSSAYIT H QRV H TGEKP

C PE C GKSFCESSPLIH H TVI H TGEKP FE C LE C GKAFNHRSYLKR H QRI H TGEKP
                                                                    Z4FPCR.H7
YV C SE C GKALTHCSTFIL H KRA H TGEKP FE C LE C GKAFSNRADLIR H FSI H TGEKP

C TG  GKSFSYVSYYTR H QRI H TGEKP FE C SE C GKAFNGNSSLIR H QRI H TGERP
                                                                    Z4FPCR.H8
YH C EE C GRAFNDNANLIR H QRI H SGDRP YY C TE C GNSFTSSSGFVI H QRI H TGEKP

C PE C GKFFNDNANLIR H QRI H SGDRP YY C TE C GNSFTSSSEFVI H QRI H TGEKP
                                                                    Z4FPCR.H9
YE C NE C GKAFVGNSPLLR H QKI H TGEKP YE C NE C GKSFSWSFHLIA H QRT H TGEKP

C PE C GKSLSSKYALVE H QRT H NGEKP YV C NV C GKSFRHKQTFVG H SRD H TGERP
                                                                    Z4FPCR.H10
YV C ME C GKSFIHSYDRIR H QSS H TRRA  YQ C SE C GKSFIYKQSLLD H HRT H TGEKP
```

FIG. 4. Zinc-finger domain sequences. The amino acid sequences of 10 of the 4 finger-containing amplification products are shown. The amino acids underlined in the Z4FPCR.H1 clone represent the primer-derived sequences present in all of the clones and, therefore, may not represent the true genomic sequence.

mids with inserts of the expected size were identified and the inserts were sequenced. The deduced amino acid sequences from 10 such inserts are shown in Fig. 4. Each insert contains primer-derived sequences at each end although in one case the region derived from the Cys-$Xaa_2$-Cys primer appears to have suffered a 3-bp deletion. The internal sequences correspond to zinc fingers as expected with all of the characteristic features including the conserved metal-ligating and hydrophobic residues and, in most cases, the H-C link sequence. One exception was observed in which a serine residue was found in a position where one of the conserved cysteines was expected. For this sequence the linker region was also one amino acid shorter than is generally observed. It is not known whether this represents a true variation or a PCR-related anomaly.

All 10 clones we have sequenced are distinct and, to our knowledge, different from published zinc-finger sequences. This observation provides further evidence concerning the large number of zinc-finger sequences within the human genome. Human cDNA libraries from various sources that have been estimated to contain 70 (11) and 100 (10) cDNAs have been reported and these appear to be significantly nonoverlapping. The fact that we have not repeatedly isolated any fragment thus far indicates that there are at least 50 species within the four-finger-encoding band we have been studying. Thus these data indicate the wide occurrence of the TFIIIA-like zinc-finger domains in the human genome. The extant data suggest that there may be several hundred zinc-finger-encoding genes. Given that the fully sequenced zinc-finger cDNAs encode on average ≈10 zinc fingers, this suggests that sequences that encode TFIIIA-like zinc-finger sequences may account for (200 zinc-finger-encoding genes × 10 zinc-finger domains per gene × 84–96 bp per domain) = $2 \times 10^5$ bp. This represents nearly 0.01% of the human genome.

Although most zinc-finger domains have two amino acids separating the cysteine residues, several important proteins have four. These include TFIIIA itself (6, 7), the general transcription factor Sp1 (17), the product of the glioma-related gene *GLI* (18), the growth factor-induced gene product zif 268 (also known as NGF-IA, egr-1, and Krox 24) (19), and a candidate for the Wilm tumor gene product (20). However, such sequences are not well represented in the sequences that have been identified from screening using low-stringency hybridization methods. To determine if our PCR method could be used to specifically amplify such sequences, a primer was designed based on the consensus sequence derived from known Cys-$Xaa_4$-Cys sequences. As shown in Fig. 5, amplification using this primer and the H-C link primer results in the generation of bands that are slightly larger than those observed from use of the Cys-$Xaa_2$-Cys and H-C link primers as expected. Thus, the PCR method appears to be capable of amplification of Cys-$Xaa_4$-Cys-containing zinc-finger sequences for cloning and characterization.

Finally, we have used our PCR method to examine the occurrence of zinc-finger-encoding sequences in various eukaryotic genomes. As shown in Fig. 6, ladders of amplification products similar to those produced from human genomic DNA were observed with most of the organisms so far examined. However, very few amplification products were observed with the yeast *Saccharomyces cerevisiae*. Zinc-finger-encoding genes have been isolated from yeast (21, 22), but these tend to contain small numbers of zinc-finger domains per gene product and the sequences of the zinc-finger domains that have been determined tend to be more different than the bulk of the sequences in the data base than those from other organisms. To confirm that the bands in these

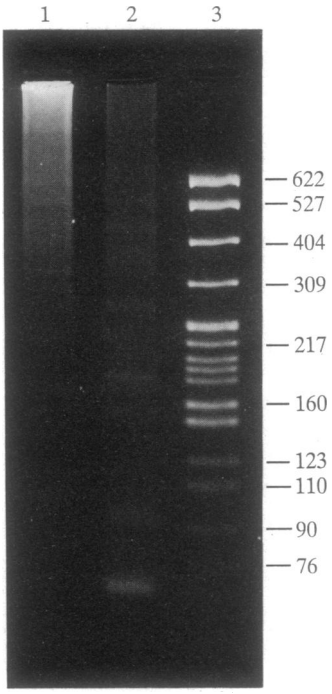


FIG. 5. Amplification of sequences encoding zinc-finger domains of the type Cys-$Xaa_4$-Cys. To determine if we would be able to specifically amplify sequence variants of the zinc-finger motif, a primer was designed to be complementary to sequences encoding the Cys-$Xaa_4$-Cys type domain. As can be seen, the fastest migrating fragment is ≈6 bp larger than the single-finger-encoding band (97 bp) in the Cys-$Xaa_2$-Cys lane as predicted. Lanes: 1, Cys-$Xaa_4$-Cys to H-C link; 2, Cys-$Xaa_2$-Cys to H-C link; 3, molecular size markers as described in Fig. 2.

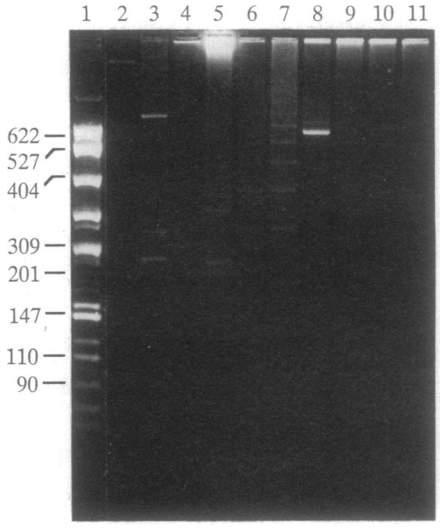


FIG. 6. Detection of zinc-finger sequences in various eukaryotic genomes. A variety of eukaryotic genomic DNAs was used as the template to amplify zinc-finger-encoding sequences using the Cys-$Xaa_2$-Cys and the H-C link primers. Ladders similar to those obtained when human genomic DNA is used as the template were obtained for many of the organisms examined. All of the reactions were carried out as described in *Materials and Methods* except that an annealing temperature of 40°C was used for the *Zea mays* incubation (lane 5). The faint bands present in the *Drosophila melanogaster* lane (lane 4) are due to a loading artifact and do not represent a lack of amplified sequences. Many of the expected sized bands are present upon amplification. Lanes: 1, molecular size markers, as described in Fig. 2; 2, *Saccharomyces cerevisiae*; 3, *Caenorhabditis elegans*; 4, *Drosophila melanogaster*; 5, *Zea mays*; 6, *Xenopus laevis*; 7, zebrafish; 8, chicken; 9, mouse; 10, rat; 11, human.

ladders do derive from the organisms tested, we have cloned fragments from *Caenorhabditis elegans* as a representative organism and shown that the inserts hybridize to specific genomic fragments on a Southern blot (data not shown). Thus, the PCR method should be useful for isolation of zinc-finger-encoding fragments from a wide variety of organisms.

In summary, we have developed methods based on PCR that allow the isolation of DNA fragments that encode TFIIIA-like zinc-finger domains. The fact that such sequences often occur in tandem arrays, even in genomic DNA, is evident from the results and this property has been used in characterizing the DNA fragments amplified. These techniques should be a useful adjunct to the low-stringency hybridization methods that have been used to collect large numbers of zinc-finger-encoding cDNA clones. Given the large size of this superfamily of genes and the important roles members of the superfamily appear to play in controlling development and other processes, the increased ability to isolate genes of this type should find considerable application.

We are grateful to Drs. Donald Brown, Ed Hedgecock, Bert Vogelstein, Nina Federoff, and Carl Spana and Doug Harrison for supplying DNA samples; Dr. John Bishop for assistance in the preparation of Fig. 6; and Dr. Victor Corces for support and encouragement. This work was supported by the National Institutes of Health Grant GM-38230 and the National Science Foundation Presidential Young Investigator Award. G.R.P. was supported by National Research Training Award 5T32-GM-07231 from the National Institutes of Health.

1. Kamb, A., Weir, M., Rudy, B., Varmus, H. & Kenyon, C. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 4372–4376.
2. Tulinsky, A., Park, C. H. & Skrzypczak-Jankun, E. (1988) *J. Mol. Biol.* **202,** 885–901.
3. Klug, A. & Rhodes, D. (1987) *Trends Biochem. Sci.* **12,** 464–469.
4. Evans, R. & Hollenberg, S. M. (1988) *Cell* **52,** 1–3.
5. Berg, J. M. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 405–421.
6. Miller, J., McLachlan, A. D. & Klug, A. (1985) *EMBO J.* **4,** 1609–1614.
7. Brown, R. S., Sander, C. & Argos, P. (1985) *FEBS Lett.* **186,** 271–274.
8. Chavrier, P., Lemaire, P., Revelant, O., Bravo, R. & Charnay, P. (1988) *Mol. Cell. Biol.* **8,** 1319–1326.
9. Knöchel, W., Pöting, A., Köster, M., El Baradi, T., Nietfeld, W., Bouwmeester, T. & Pieler, T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6097–6100.
10. Bellefroid, E. J., Lecocq, P. J., Benhida, A., Poncelet, D. A., Belayew, A. & Martial, J. A. (1989) *DNA* **8,** 377–387.
11. Thiesen, H.-J. (1990) *New Biologist* **2,** 363–374.
12. Chowdhury, K., Deutsch, U. & Gruss, P. (1987) *Cell* **48,** 771–778.
13. Schuh, R., Aicher, W., Gaul, U., Côté, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R. & Jäckle, H. (1986) *Cell* **47,** 1025–1032.
14. Dente, L., Cesareni, G. & Cortese, R. (1983) *Nucleic Acids Res.* **11,** 1645–1656.
15. Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J. & Struhl, K., eds. (1987) *Current Protocols in Molecular Biology* (Wiley, New York).
16. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Ehrlich, H. A. (1988) *Science* **239,** 487–491.
17. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987) *Cell* **51,** 1079–1090.
18. Kinzler, K. W., Ruppert, J. M., Bigner, S. H. & Vogelstein, B. (1988) *Nature (London)* **332,** 371–374.
19. Christy, B., Lau, L. F. & Nathans, D. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7857–7861.
20. Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Karl, A., Yeger, D. A., Lewis, W. H., Jones, C. & Housman, D. E. (1990) *Cell* **60,** 509–520.
21. Hartshorne, T. A., Blumberg, H. & Young, E. T. (1986) *Nature (London)* **320,** 283–287.
22. Stillman, D. J., Bankier, A. T., Seddon, A., Groenhout, E. G. & Nasmyth, K. A. (1988) *EMBO J.* **7,** 485–494.