Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 9563–9567, November 1991
Biochemistry

Characterization and mapping of human genes encoding zinc finger proteins

(transcription/chromosome/sequence-tagged site)

P. BRAY*, P. LICHTER†‡, H.-J. THIESEN§, D. C. WARD†, AND I. B. DAWID*

*Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892; †Department of Genetics, Yale School of Medicine, New Haven, CT 06520; ‡Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-6900 Heidelberg, Federal Republic of Germany; §Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland

Contributed by I. B. Dawid, August 2, 1991

ABSTRACT **The zinc finger motif, exemplified by a segment of the *Drosophila* gap gene Krüppel, is a nucleic acid-binding domain present in many transcription factors. To investigate the gene family encoding this motif in the human genome, a placental genomic library was screened at moderate stringency with a degenerate oligodeoxynucleotide probe designed to hybridize to the His/Cys (H/C) link region between adjoining zinc fingers. Over 200 phage clones were obtained and are being sorted into groups by partial sequencing, cross-hybridization with oligodeoxynucleotide probes, and PCR amplification. Further, the genomic clones were cross-hybridized with a set of 30 zinc finger-encoding cDNAs (Kox1–Kox30) isolated from a human T-cell cDNA library. Four cDNAs (Kox4, Kox7, Kox12, and Kox15) were identified that match one or more genomic clones; these matches were confirmed by nucleotide sequence analysis. One or more clones from each locus were mapped onto human metaphase chromosomes by chromosomal *in situ* suppression hybridization with fluorescent probe detection. We mapped ZNF7/Kox4 to chromosome 8qter, ZNF19/Kox12 to 16q22, ZNF22/Kox15 to 10q11, and ZNF44/Kox7 to 16p11. The results of these analyses support the conclusion that the human genome contains many, probably several hundred, zinc finger genes with consensus H/C link regions.**

The zinc finger DNA-binding domain was first identified in the *Xenopus* RNA polymerase III transcription factor TFIIIA (1, 2). Zinc fingers contain ≈30 amino acids with a consensus sequence of $CX_{2,4}CX_3FX_5LX_2HX_{3,4}H$, where X indicates amino acids that are less well conserved. Finger folding is organized around a tetrahedrally coordinated zinc ion bound by the conserved cysteine (C) and histidine (H) residues (2, 3). Fingers occur in tandem arrays with a minimum of two consecutive units; TFIIIA has nine consecutive zinc finger domains, and both shorter and longer arrays have been reported (4–10). Most well-studied finger proteins have been characterized as DNA-binding proteins. Several developmental control genes (e.g., Krüppel; ref. 9), transcription factor genes (e.g., Sp1; ref. 7), and genes correlated with diseases (e.g., Wilms tumor; ref. 11) contain zinc finger domains, attesting to the importance of this structure in the regulation of gene expression.

The *Drosophila* gap gene Krüppel is the prototype of a large subfamily of zinc finger genes in which the amino acids joining adjacent fingers, TGEKPYE/K (the H/C link), are more highly conserved than those within the fingers except for the metal-coordinating residues (12). Oligodeoxynucleotides and cDNA fragments representing the H/C link consensus sequence have been used to isolate C_2H_2 zinc finger genes from *Xenopus* (8, 13) and human (4) cDNA libraries and from mouse genomic libraries (6). Unlike zinc finger genes cloned by binding of known regulatory nucleotide sequences, zinc finger clones isolated by sequence homology often contain many fingers and highly conserved H/C link regions. Among multifinger cDNAs at least two associated motifs of unknown function have been identified and named FAX (finger-associated box; ref. 14) and KRAB (Krüppel-associated box; ref. 15).

With the aim to survey the number and chromosomal distribution of zinc finger genes in the human genome, we initiated a screen with a probe corresponding to the conserved H/C link region. We searched a genomic library rather than cDNA libraries to avoid any bias due to differential expression of individual zinc finger genes. This paper describes our general approach to the characterization of zinc finger genes and reports partial sequence information and mapping data for four genes[¶] that correspond to clones previously isolated by one of us from a T-cell cDNA library (10).

MATERIALS AND METHODS

Clone Isolation. A Charon 28 library containing human placental DNA inserts obtained by partial *Sau*3A1 digestion was generously provided by P. Leder (Harvard Medical School). Approximately 5×10^5 phage were screened by hybridization in a solution containing 5× standard saline citrate (SSC), 5× Denhardt's solution, 50 mM sodium phosphate (pH 6.8), 1 mM sodium pyrophosphate, 50 μg of yeast tRNA per ml, 10% (vol/vol) formamide, and ^{32}P-labeled oligodeoxynucleotide probe (1.5×10^6 cpm, 175 fmol/ml; labeled at the 5′ end by using polynucleotide kinase). The probe was based on aligned nucleotide sequences of a set of C_2H_2 fingers and incorporates only some of the degeneracy of the sequence set. The probe sequence and the encoded amino acid sequence are as follows.

G AA^C_G CAC C^A_TG ^C_AGG A^{CG}_{TT} CAC ACG GG^G_A GAG AAG CCC TAC CA

K/N H L/Q R I/M/T H T G E K P Y

Filters were hybridized for 16 hr at 42°C, washed in 0.2× SSC/0.1% SDS three times at 23°C for 10 min per wash and once at 39°C for 2 min, and then subjected to autoradiography.

Sequencing. Phage DNA was cleaved by several restriction enzymes, and fragments that hybridized with the oligodeoxynucleotide probe shown above were subcloned into M13mp18 or pBluescript. Nucleotide sequences were obtained by the dideoxynucleotide chain-termination method (16). Sequences were analyzed with the Genetics Computer Group sequence-analysis package (17).

Dot Blot Hybridization and PCR Amplification. Phage DNA was prepared from each clone and dot blots of all phage DNAs were hybridized with either cDNA or oligonucleotide

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: DAPI, 4′,6-diamidino-2-phenylindole; FITC, fluorescein isothiocyanate; T_m, melting temperature.

[¶]The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M77170–M77173).

probes at high stringency. Dot blots were prehybridized (1 hr) and hybridized (24 hr) with ^{32}P-labeled cDNA in 0.5 M sodium phosphate, pH 7.2/7% SDS/1 mM EDTA at 67–69°C. Dot blots were washed in 40 mM sodium phosphate, pH 7.2/0.1% SDS/1 mM EDTA at 67–69°C twice for 30 min and exposed to Kodak XAR5 film for 1 hr to 24 hr in the presence of an intensifying screen. For oligonucleotide hybridization, blots were incubated in 6× SSC/5× Denhardt's solution/0.1% SDS with ^{32}P-labeled oligonucleotide probe (1.0×10^6 cpm/ml, 120 fmol/ml) at 5°C below the melting temperature (T_m − 5°C) for 4 hr and then washed in 6× SSC three times at 23°C for 5 min and once at T_m − 10°C for 2 min (see legend to Fig. 2 for description of probes and T_m). PCR reactions were carried out using the AmpliTaq reagent kit (Cetus) for 25 cycles of 95°C, 1 min; 60°C, 1 min; 72°C, 1 min. One-half microliter of phage supernatant was used as a source of DNA template in each 20-μl reaction mixture. One reaction was carried out with a complete reaction mixture including primers but without DNA template.

Chromosome Mapping. Human metaphase chromosome spreads were prepared from cultured lymphocytes by standard methods of Colcemid arrest, hypotonic treatment, and methanol/acetic acid fixation. Phage DNA probes were labeled with biotin or digoxigenin by nick-translation and hybridized *in situ* under suppression hybridization conditions as reported elsewhere (18, 19). Biotinylated DNAs were used in single probe hybridizations and detected via avidin-conjugated fluorescein isothiocyanate (FITC). To visualize chromosomal bands, chromosomes either were counterstained with propidium iodide and simultaneously banded with 4′,6-diamidino-2-phenylindole (DAPI) dihydrochloride or were cohybridized with cloned *Alu* DNA sequences (19, 20). For confirmation of chromosome assignment and for detailed regional mapping, the biotinylated probes were cohybridized with chromosome-specific DNA probes or probe sets labeled with digoxigenin. The following probes were used for cohybridization experiments: the DNA library pBS8, derived from sorted and amplified human chromosome 8 DNA sequences (21) and generously provided by Joe Gray (Lawrence Livermore National Laboratory), the alphoid probe *D10Z1*, labeling the centromere of chromosome 10 (22), kindly provided by Thomas Cremer (University of Heidelberg); and the alphoid probe pSE16 (*D16Z2*), labeling the centromeric region of human chromosome 16 (23), kindly supplied by Huntington Willard (Stanford University). For each probe, a minimum of 35 metaphase spreads were evaluated and at least 80% of the metaphases showed specific signals on all four chromatids.

RESULTS

Strategy. Our aim is to characterize and map a large set of genomic clones encoding putative zinc finger proteins. In the present paper we describe the general approach we have used, and present specific results on four genes of the set.

On initial screening of a genomic library with a degenerate oligodeoxynucleotide probe (see *Materials and Methods*) we picked 208 clones from about 600 positives among 5×10^5 plaques. Within the set of 208 clones we expected multiple clones representing overlapping regions of the same locus, in addition to duplicate isolations of identical clones. The first task was thus to sort clones into genomic loci. At the same time we wished to determine which of the clones in the set of 208 correspond to members of a set of 30 zinc finger protein-encoding cDNAs (Kox1–Kox30) previously isolated from a T-cell library (10).

For sorting of genomic clones we prepared DNA from all phages and hybridized sets of Southern and dot blots with each of the 30 cDNA clones. Tentative matches were then checked by partial sequence analysis and by additional hybridization and PCR amplification. To obtain sequence information, we digested selected genomic clones with restriction enzymes expected to yield fragments of several

```
Clone 38 (ZNF7)

ATCTCTCGCCTGAGTCAGCATCAGCTGATTCACACTGGAGAGAAGCCTTATAAATGCAAC
 I  S  R  L  S  Q [H] Q  L  I [H] T  G  E  K  P  Y  K [C] N

AAGTGTACAAAAGCCTTTGGTTGTAGTTCACGGCTTATTCGCCATCAGAGAACTCACACT
 K [C] T  K  A  F  G  C  S  S  R  L  I  R [H] Q  R  T [H] T

GGAGAAAAACCATTTAAATGTGATGAGTGTGGCAAAGGCTTTGTTCAGGGCTCACACCTT
 G  E  K  P  F  K [C] D  E [C] G  K  G  F  V  Q  G  S  H  L

ATTCAGCATCAGCGAATCCACACTGGAGAGAAACCCTATGTGTGTAATGACTGTGGAAAA
 I  Q [H] Q  R  I [H] T  G  E  K  P  Y  V [C] N  D [C] G  K

GCCTTCAGTCAGAGTTCCAGCCTTATTTACCATCAGAGAATCCATAAAGGAGAGAAGCCC
 A  F  S  Q  S  S  S  L  I  Y [H] Q  R  I [H] K  G  E  K  P

TACGAATGCCTCCAATGCGGAAAAGCCTTCAGTATGAGCACACAGCTTACAATACATCAA
 Y  E [C] L  Q [C] G  K  A  F  S  M  S  T  Q  L  T  I [H] Q

AGGGTTCACACTGGAGAGAGGCCCTATAAATGTAATGAATGTGGGAAAGCCTTCAGTCAA
 R  V [H] T  G  E  R  P  Y  K [C] N  E [C] G  K  A  F  S  Q

AACTCAACCCTTTTCCAACACCAGATAATTCATGCAGGGGTGAAGCCCTATGAGTGCAGT
 N  S  T  L  F  Q [H] Q  I  I [H] A  G  V  K  P  Y  E [C] S

GAGTGTGGAAAAGCC
 E [C] G  K  A


Clone 84 (Kox7)

CCCTATGAATGTAAGCAATGTGGGAAATTGTTATCTCATCGCTCAAGCTTTCGAAGACAC
 P  Y  E [C] K  Q [C] G  K  L  L  S  H  R  S  S  F  R  R [H]

ATGATGGCACACACTGGAGATGGCCCTCATAAATGCACAGTATGTGGGAAAGCCTTTGAT
 M  M  A [H] T  G  D  G  P  H  K [C] T  V [C] G  K  A  F  D

TCTCCTAGTGTATTTCAAAGACATGAAAGGACTCACACTGGAGAGAAACCCTATGAATGC
 S  P  S  V  F  Q  R [H] E  R  T [H] T  G  E  K  P  Y  E [C]


Clone 114 (ZNF7)                    ..

CACACTGGGGAGAAACCCTATGAGTGTAATGAGTGTGGCAAAGCTTTTGTTGGTAATTCA
[H] T  G  E  K  P  Y  E [C] N  E [C] G  K  A  F  V  G  N  S

CCCCTACTTCGGCATCAGAAAATCCACACTGGAGAGAAACCCTATGAGTGTAATGAGTGT
 P  L  L  R [H] Q  K  I [H] T  G  E  K  P  Y  E [C] N  E [C]

GGCAAAAGCTTTGGAAGGACTTCCCATCTAAGCCAACATCAGCGTATTCACACAGGGGAA
 G  K  S  F  G  R  T  S  H  L  S  Q [H] Q  R  I [H] T  G  E

AAGCCTTATTCTTGTAAAGTATGTGGACAAGCCTTCAATTTTCATACAAAACTAACTCGG
 K  P  Y  S [C] K  V [C] G  Q  A  F  N  F  H  T  K  L  T  R

CACCAGAGAATT
[H] Q  R  I


Clone 90 (ZNF22)

ATTCATACTGGAGAAAAACCCTATCAGTGTGATGAGTGTGGCCGGTGTTTCAGCCAGAGC
 I [H] T  G  E  K  P  Y  Q [C] D  E [C] G  R  C  F  S  Q  S

TCCCACCTTATTCAACATCAGAGAACCCACACTGGGGAGAAACCCTACCAGTGCAGTGAA
 S  H  L  I  Q [H] Q  R  T [H] T  G  E  K  P  Y  Q [C] S  E

TGTGGCAAATGTTTCAGTCAGAGCTCTCATCTGAGGCAGCACATGAAGGTGCATAAAGAA
[C] G  K  C  F  S  Q  S  S  H  L  R  Q [H] M  K  V [H] K  E

GAGAAGCCTCGTAAAACCCGGGGCAAAAATATCAGGGTGAAGACTCACTTACCCTATTGG
 E  K  P  R  K  T  R  G  K  N  I  R  V  K  T  H  L  P  Y  W

AAGCTGGTACAGGAA
 K  L  V  Q  E
```

FIG. 1. Nucleotide and amino acid sequences of portions of the zinc finger regions of four genomic clones. The overlapping regions of these clones are identical in nucleotide sequence with the corresponding cDNAs listed in parentheses. The cDNA sequence for ZNF7 is from ref. 24; amino acid sequences for the other three cDNAs were shown in ref. 10. The conserved cysteine and histidine residues characteristic of the C_2H_2 fingers are boxed.

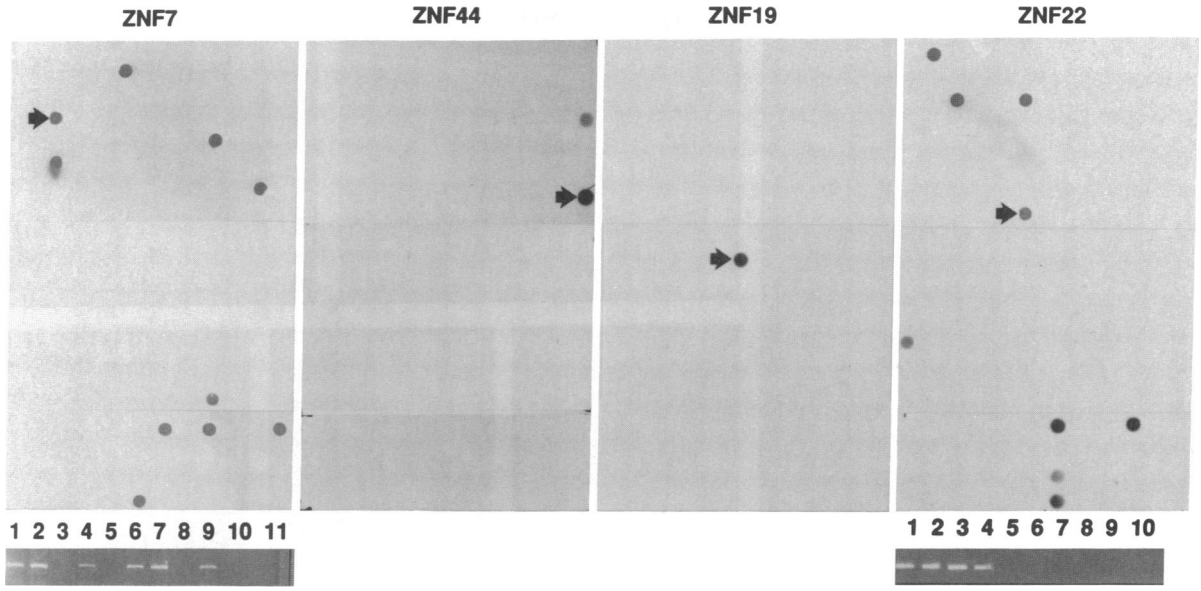

FIG. 2. The upper panels show hybridization of ^{32}P-labeled locus-specific oligonucleotides to dot blot grids of the 208 phage DNAs. An arrow points to the clone from which finger-encoding nucleotide sequence was determined (see Fig. 1). For each probe the calculated T_m is listed; this value was taken as the basis for establishing washing conditions (see *Materials and Methods*). ZNF7/Kox4, 5′-CTCTCGCCTGAGTCAGCATCAGCTG-3′ (T_m = 82°C). ZNF44/Kox7, 5′-ATGAAGGAACTCACACTC-3′ (T_m = 60°C); this sequence is derived from the corresponding cDNA clone (10). ZNF19/Kox12, 5′-TCTGGTGCCGAGTTAGTT-3′ (T_m = 54°C). ZNF22/Kox15, 5′-GAAGACTCACTTACCCTC-3′ (T_m = 54°C). Below the dot blot grids For ZNF7 and ZNF22 are shown PCR analyses of all clones that hybridized with the respective oligonucleotides. For ZNF7, the hybridization primer and the primer 5′-CAGTCTTCACACTGATGCA-3′ derived from the cDNA sequence (24) led to amplification of the predicted 641-base-pair fragment from six genomic clones which are therefore listed as members of the ZNF7 locus in Table 1. Four clones and the control (lane 11, no DNA) yielded no band of this size. For ZNF22, two primers based on the sequence of clone 30, 5′-CCTATACTTAACGGAGGCCAGCCAC-3′ and 5′-CGGTGTTTCAGCCAGAGCTCCCAC-3′, yielded the predicted 241-base-pair fragment from four clones; these are listed under ZNF22 in Table 1. Five clones and the control (lane 10) did not yield a band of this size.

hundred nucleotides. These were subcloned, and plaques or colonies were hybridized with the original zinc finger consensus oligodeoxynucleotide (*Materials and Methods*). Positive subclones were isolated and sequenced, yielding sequences that largely though not exclusively corresponded to finger domains. These sequences were used to check the tentative assignments of matches with cDNA clones and to prepare locus-specific oligodeoxynucleotide probes for additional hybridizations.

For further sorting, aliquots of DNA from the 208 phages were spotted on replicate dot blot grids; these grids were then hybridized with the locus-specific oligodeoxynucleotide probes. Positive phages were tentatively grouped together. To test the groupings, primer pairs were derived from the sequence of the prototype clone of the group and were used to amplify fragments from each presumed member of the group. Those phages that yielded PCR fragments of the size predicted from the sequence were considered to have been derived from the same locus. Other phages that hybridized with the probe were classified as having been derived from related genes. Within each group we determined the minimal number of different (i.e., overlapping) phages by examining restriction digests; the remaining clones were considered as probable duplicate isolates of the same phage. Finally, representative members of different groups (or loci) were mapped onto metaphase chromosomes.

Matches Between cDNAs and Groups of Genomic Clones. Comparison of the set of 30 T-cell cDNA clones (10) with the set of 208 genomic clones yielded only four matches by high-stringency hybridization. These matches were confirmed by partial sequencing of one genomic clone from each group (Fig. 1).

Table 1. Classification of genomic clones

Locus	cDNA	Genomic clones
ZNF7	Kox4	17, 38, 62, 189, 199, 204
ZNF44	Kox7	48, 84
ZNF19	Kox12	114
ZNF22	Kox15	2, 27, 30, 90

Next, oligodeoxynucleotides specific for prototype genomic clones representing each locus were synthesized and used in dot blot hybridizations with complete grids of the 208 phage DNAs (Fig. 2). Kox7 (ZNF44) corresponded to two identical genomic clones and Kox12 (ZNF19) to a single clone (Table 1); in contrast, Kox4 (ZNF7)- and Kox15 (ZNF22)-specific probes hybridized to multiple clones. To check whether all these clones were indeed derived from the ZNF7 or ZNF22 locus, we prepared PCR primer sets specific for each locus and used them to attempt amplification from each of the putative members of the two groups. In the case of the ZNF7 group, six genomic clones acted as templates for amplification (Fig. 1) and are thus listed as members of the group in Table 1; four other cross-hybridizing clones did not produce the appropriate PCR products and are considered to represent related genes. In the ZNF22 group four clones passed the PCR test; of these, clones 2, 27, and 90 are identical, whereas clone 30 is overlapping.

The dot blot hybridization and PCR amplification experiments of Fig. 2 thus classify 13 of 208 zinc finger-encoding clones into four loci. The remaining 195 clones represent an unknown number of additional loci.

Comparison of Finger Sequences. Each of the four clones considered here encodes tandem finger domains commonly found in C_2H_2 zinc finger proteins (4–10). The sequences obtained are only partial, and it is likely that additional finger domains are located beyond the ends of the available sequence. Most of the H/C link regions in these genes are close to the consensus for this subfamily, TGEKPY (12).

Chromosomal Mapping. *In situ* suppression hybridization with biotinylated whole phage DNA yielded specific signals well above background fluorescence that greatly reduced the

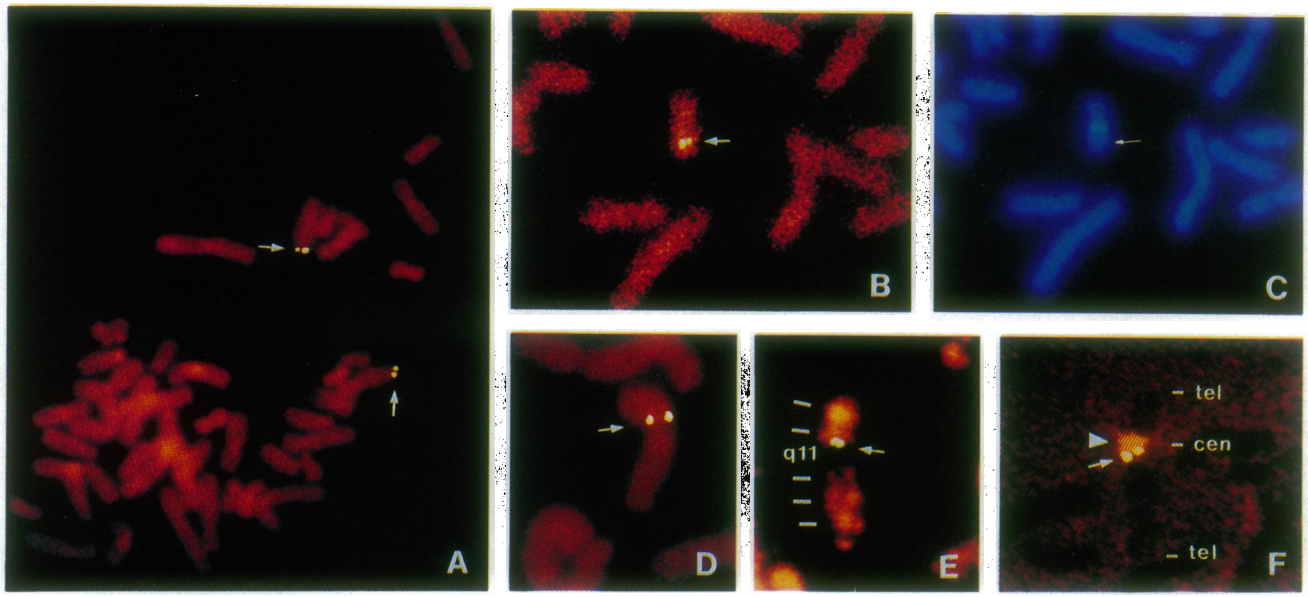

FIG. 3. Mapping of four zinc finger genes to human chromosomes by *in situ* suppression hybridization. The genomic clones 38 (*A*), 114 (*B*), 84 (*D* and *E*), and 27 (*F*) were biotinylated, hybridized, and detected via avidin-conjugated FITC. (*A*) A metaphase spread of propidium iodide-counterstained chromosomes after hybridization with clone 38 (ZNF7/Kox4) probe. Highly specific signals are seen on both chromatids of both homologues of 8qter (arrows). (*B*) FITC signal (arrow) of clone 114 (ZNF19/Kox12) on propidium iodide-counterstained chromosome 16q. (*C*) DAPI staining of chromosomes seen in *B*. Band 16q22 is indicated (arrow). (*D*) ZNF44/Kox7 probe hybridizing to proximal part of 16p (arrow), seen on a propidium iodide-counterstained chromosome. (*E*) For fine mapping, *Alu* banding was performed, resulting in an R-banding-like pattern. Since there is no substantial amount of *Alu* DNA in the heterochromatic region of 16q11, this region appears as a gap on the banded chromosome (see label). Bars indicate the R bands delineated by the *Alu* probe in the order (from the top): 16p13, p11, q13, q22, and q24. Note the localization of the probe to the distal portion of 16p11. (*F*) Mapping of the ZNF22/Kox15 gene (arrow) to 10cen just distal from the cohybridized alphoid probe *D10Z1* (arrowhead), which was labeled with digoxigenin and detected via rhodamine. The positions of the telomeres (tel) and the centromere (cen) are indicated. The dark outline of the chromosomes was achieved by digital filtering of the image. *A*, *B*, *D*, *E*, and *F* are digitized images obtained by using a confocal laser scanning microscope (Zeiss LSM10) as described (19). The picture in *C* was taken by using a conventional epifluorescence microscope (Zeiss Axioplan).

need for statistical analysis of the data (Fig. 3). Cohybridization with chromosome-specific probes was used for verification of chromosomal position. The genomic clones used for the mapping study and the cohybridization probes are listed in Table 2. In order to give cytological mapping coordinates, chromosomes were banded with DAPI and *Alu* probe as indicated in Table 2. Three of the four genes were localized at distinct morphological sites of the chromosomes: at telomeric (ZNF7/Kox4) and centromeric (ZNF44/Kox7 and ZNF22/Kox15) regions. Both phage DNAs containing sequences of ZNF7 mapped extremely distal on the long arm of chromosome 8 (Fig. 3*A*)—i.e., in the distal portion of band 8q24. The localization of ZNF7 to 8qter is in agreement with previous ZNF7 cDNA mapping by isotopic *in situ* hybridization (24). The probe for ZNF44/Kox7 mapped to the centromeric region of chromosome 16 (Fig. 3*D*). To fine map the gene relative to the centromere, cohybridization with an alphoid probe specifically delineating 16cen (data not shown) and *Alu* banding (Fig. 3*E*) were performed. Both experiments revealed the regional localization of ZNF44 just adjacent to the alphoid sequences on 16p. *Alu* banding, which results in an R-banding-like pattern (25), revealed the mapping of ZNF44 to the proximal part of band 16p11 (Fig. 3*E*). ZNF19/Kox12 is localized on the same chromosome but on distal 16q (Fig. 3 *B* and *C*). Extrapolating the mapping coordinates obtained from chromosome length measurements (19) suggested a localization in 16q22. This was confirmed by *Alu* banding (data not shown). The mapping of the two probes for ZNF22/Kox15 to 10cen was further investigated by cohybridization with a probe specifically delineating the alphoid DNA on human chromosome 10 (Fig. 3*F*). To determine the precise topological relation of the phage probe to 10cen, carefully adjusted digitized images were generated from both signals. The overlay of the two images then showed that the ZNF22 gene is located just distal to the area delineated by the alphoid sequence probe (Fig. 3*F*). Occasionally, the FITC fluorescence spots of the phage probe were even overlapping with the rhodamine-stained centromeric area. These findings suggest that the cytological position of ZNF22 is 10q11. Kox15 has been mapped to 10cen–q24 (26).

DISCUSSION

In this paper we discuss an approach to the characterization of a set of human DNA sequences isolated on the basis of homology to the conserved motif of the Krüppel class of zinc finger protein-encoding genes. Since the aim is to sort many

Table 2. Localization of genomic clones

Locus	Genomic clone	Map position	Cohybridization	Banding
ZNF7	17	8qter	Not done	DAPI, *Alu*
	38	8qter	pBS8 (Chr library)	DAPI
ZNF44	84	16p11	pSE16 (alphoid)	DAPI, *Alu*
ZNF19	114	16q22	pSE16 (alphoid)	DAPI, *Alu*
ZNF22	27	10cen (q11)	*D10Z1* (alphoid)	DAPI
	30	10cen (q11)	*D10Z1* (alphoid)	DAPI

Chr, chromosome.

genes into groups, we used a combination of partial sequencing, cross-hybridization with oligonucleotides, and PCR amplification as tools of classification. Finally, clones representing each group, or gene, were mapped onto the human chromosome complement. The use of PCR with specific oligodeoxynucleotide pairs as a main tool of characterization provides the advantages of the proposed "sequence-tagged site" strategy (27) in reducing error accumulation and providing convenient means to check results, both of which are needed in dealing with large numbers of clones.

When isolating genomic sequences by homology to a conserved motif, one must consider whether each resulting isolate corresponds to an active gene or a pseudogene. It is likely that the four loci discussed in this paper represent active genes, because of their correspondence to cDNA clones isolated previously. Each of the four sequences maps to a single distinct genomic site, arguing against the possibility that any of these sequences represents both an active gene and a pseudogene.

In screening 2.5 genome equivalents we noted about 600 positive phages, 238 of which were picked, resulting in 208 phages that remained positive after rescreening. This observation suggests that there are about 200 phages per genome equivalent that are capable of hybridizing with the degenerate oligodeoxynucleotide representing the H/C link region of the Krüppel class of zinc finger genes. The four loci discussed in this paper are represented by an average of 3.25 phages in our 208-clone set; analyses of other clones in the set suggest that this average is representative, and thus the clone set contains between 50 and 100 zinc finger genes. In comparing this set to the set of 30 zinc finger cDNAs isolated previously from a T-cell library (10), only 4 matches were found. Making assumptions about randomness of sampling, one may again estimate that the genome contains several hundred zinc finger genes. This impression is strengthened by the fact that only 1 match between the set of genomic clones and other published cDNAs was found: ZNF7 has been reported by Lania *et al.* (24) and is also a member of the T-cell cDNA set (10). Additional sets of clones from human cDNA libraries have been reported (4, 28, 29), leading to similar conclusions about the size of the zinc finger gene family. Likewise, extensive analysis of this gene family in *Xenopus laevis* has led to the characterization of at least 100 zinc finger genes in the frog genome (8). It is thus likely that vertebrate animals in general contain large sets of such genes.

Most zinc finger proteins that have been characterized functionally are DNA-binding proteins and transcription factors (e.g., Sp1; ref. 7). However, the original zinc finger protein, TFIIIA, functions as an RNA-binding protein as well as a transcription factor (30, 31). It is therefore possible that some of the proteins encoded by the genes discussed here are RNA-binding, but it remains likely that the majority of zinc finger proteins are transcription factors. Thus, the set of transcription factors derived from just a single structural family may number in the hundreds.

1. Brown, R. S., Sander, C. & Argos, P. (1985) *FEBS Lett.* **186,** 271–274.
2. Miller, J., McLachlan, A. D. & Klug, A. (1985) *EMBO J.* **4,** 1609–1614.
3. Klug, A. & Rhodes, D. (1987) *Trends Biochem. Sci.* **12,** 464–469.
4. Bellefroid, E., Lecocq, P. J., Benhida, A., Poncelet, D. A., Belayew, F. & Martial, J. A. (1989) *DNA* **8,** 377–387.
5. Chavrier, P., Lemaire, P., Revelant, O., Bravo, R. & Charnay, P. (1988) *Mol. Cell. Biol.* **8,** 1319–1326.
6. Chowdhury, K., Deutsch, U. & Gruss, P. (1987) *Cell* **48,** 771–778.
7. Kadonga, J. T., Carner, K. R., Masiarz, K. R. & Tjian, R. (1987) *Cell* **51,** 1079–1090.
8. Nietfeld, W., El Baradi, T., Mentzel, H., Pieler, T., Köster, M., Pöting, A. & Knöchel, W. (1989) *J. Mol. Biol.* **208,** 639–659.
9. Rosenberg, U. B., Schröder, C., Priess, A., Kienlin, A., Côté, S., Riede, I. & Jäckle, H. (1986) *Nature (London)* **319,** 336–339.
10. Thiesen, H.-J. (1990) *New Biol.* **2,** 363–374.
11. Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C. & Housman, D. E. (1990) *Cell* **60,** 509–520.
12. 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.
13. Ruiz i Altaba, A., Perry-O'Keefe, H. & Melton, D. A. (1987) *EMBO J.* **6,** 3065–3070.
14. Knöchel, W., Pöting, A., Köster, M., El Baradi, T., Nietfeld, W., Bouwmeester, T. & Pieler, T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6097–6100.
15. Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O. & Martial, J. A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3608–3612.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
17. Devereux, J., Haerberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
18. Lichter, P., Cremer, T., Borden, J., Manuelides, L. & Ward, D. C. (1988) *Hum. Genet.* **80,** 224–234.
19. Lichter, P., Tang, C. C., Call, K., Hermanson, G., Evans, G. A., Housman, D. & Ward, D. C. (1990) *Science* **247,** 64–69.
20. Lichter, P., Boyle, A., Cremer, T. & Ward, D. C. (1991) *Genet. Anal. Tech. Appl.* **8,** 24–35.
21. Fuscoe, J. C., Collins, C. C., Pinkel, D. & Gray, J. W. (1989) *Genomics* **5,** 100–109.
22. Devilee, P., Thierry, R. F., Kievits, T., Kolluri, R., Hopman, A. H. N., Willard, H. F., Pearson, P. L. & Cornelisse, C. J. (1988) *Cancer Res.* **48,** 5825–5835.
23. Greig, G. M., England, S. B., Bedford, H. M. & Willard, H. F. (1989) *Am. J. Hum. Genet.* **45,** 862–872.
24. Lania, L., Donti, E., Pannuti, A., Pascucci, A., Pengue, G., Feliciello, I., LaMantia, G., Langrancone, L. & Pelicci, P. G. (1990) *Genomics* **6,** 333–340.
25. Manuelidis, L. & Ward, D. C. (1984) *Chromosoma* **91,** 28–38.
26. Huebner, K., Druck, T., Croce, C. M. & Thiesen, H.-J. (1991) *Am. J. Hum. Genet.* **48,** 726–740.
27. Olson, M., Hood, L., Cantor, C. & Botstein, D. (1989) *Science* **245,** 1434–1435.
28. Pannuti, A., Lanfrancone, L., Pascucci, A., Pelicci, P.-G., LaMantia, G. & Lania, L. (1988) *Nucleic Acids Res.* **16,** 4227–4237.
29. Ruppert, J. M., Kinzler, K. W., Wong, A. J., Bigner, S. H., Kao, F.-T., Law, M. L., Seuanez, H. N., O'Brien, S. J. & Vogelstein, B. (1988) *Mol. Cell. Biol.* **8,** 3104–3113.
30. Pelham, H. R. B. & Brown, D. D. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 4170–4174.
31. Honda, B. M. & Roeder, R. G. (1980) *Cell* **22,** 119–126.