Proc. Nati. Acad. Sci. USA Vol. 88, pp. 7923-7927, September 1991 **Genetics** 

## A cluster of related zinc finger protein genes is deleted in the mouse embryonic lethal mutation  $t^{wls}$

( $t$  complex/partial  $t$  haplotype/gene cluster/deletion/gastrulation)

PHILIP H. CROSSLEY AND PETER F. R. LITTLE<sup>\*</sup>

Department of Biochemistry, Imperial College of Science Technology and Medicine, London SW7 2AZ, United Kingdom

Communicated by M. F. Lyon, June 4, 1991 (received for review January 15, 1991)

ABSTRACT We report that <sup>a</sup> number of related zinc finger protein genes are closely linked on mouse chromosome 17. At least four of these genes are transcribed in the 8.5-day postcoitum embryo and are deleted in the  $t$  complex early acting embryonic lethal mutation  $t^{w18}$ . We have evidence that additional finger protein genes are located in this region. These findings demonstrate that related finger protein genes can be clustered in the murine genome and identify genes that may be considered as candidates for the  $t^{w18}$  mutation.

The zinc finger motif is a common nucleic acid binding protein structure that can be represented as a consensus sequence of  $CX_2CX_3FX_5LX_3HX_3H$  (1, 2). Many proteins contain multiple fingers, up to 36 (3), that are often, but not invariably, "linked" by the sequence TGEKPYX following the terminal histidine. Originally identified in the Xenopus laevis transcription factor TFIIIa (4, 5), the motif has been found in other transcription factors, Spl (6), ADR1 (7), and SWI5 (8), in genes known to be critical to complex developmental processes in Drosophila melanogaster, including Krüppel  $(Kr)$  (9), snail (10), hunchback (11), and glass (12), and in genes associated with cell proliferation and tumorigenesis, evi-J and gli (13, 14). suvar, a suppressor of position effect variegation in Drosophila (15), and p43 (16), a protein that binds 5S RNA but not the 5S genes in Xenopus, show that not all finger proteins have a role in control of transcription.

Hybridization of genomic and cDNA libraries with the Kr gene and with oligonucleotide probes against the "link" regions of the finger arrays suggests that finger protein genes constitute a super family in vertebrates with perhaps as many as 300-700 members (ref. 17; unpublished data).

In this paper we identify a cluster of finger protein genes on mouse chromosome 17 that are contained within the distal inversion of the  $t$  complex and that are deleted in the partial t haplotype  $t^{w/8}$ <sup>+</sup>. The biological properties of the t complex (reviewed in ref. 18) are derived from the inversion of at least two regions of DNA on the  $t$  chromosome compared to wild-type chromosome 17 (19), resulting in suppression of recombination over the whole of the region. More than 20 complementing recessive embryonic lethal genes (20) have been mapped to the region, and animals may be homozygous only for *t* chromosomes that are of different lethal complementation groups; homozygosity for a given  $t$  lethal results in embryonic death at a characteristic age.

Partial *t* haplotypes, generated by rare recombination events across *t* complex inversions, contain only part of the full t chromosome DNA. One such partial haplotype is  $t^{w18}$ which has been shown by Bucan et al. (21) to be the product of recombination within the distal inversion, resulting in loss of four DNA markers and duplication of others. The lethal

associated with  $t^{w18}$  maps genetically to the region of the deletion (21) and is a member of the  $tcl-4$  complementation group of t lethals; this group includes  $t^{w18}$ ,  $t^4$ , and  $t^{ksl}$ . Homozygous embryos have deranged mesoderm formation and some evidence for an earlier regionalized disruption of cell division has been advanced by Snow and Bennett (22). Death normally occurs 7.5-10 days postcoitum.

## MATERIALS AND METHODS

Screening for Zinc Finger Protein Motifs. Finger protein cDNA phages were isolated from a Agt1O library made from L1210 mRNA by hybridization with <sup>a</sup> 144-fold redundant oligonucleotide (the "finger" oligonucleotide), CA(C/ T)AC(A/T)GG(A/T/G)GA(A/G)AA(A/G)CC(T/C/A)TA, which, when translated in the first reading frame, would encode the protein sequence His-Thr-Gly-Glu-Lys-Pro-Tyr, the common motif that links consecutive fingers. The finger oligonucleotide was radiolabeled with  $32P$  to a specific activity of  $1-3 \times 10^9$  cpm/ $\mu$ g using polynucleotide kinase and standard procedures (23). Prehybridization and hybridization of DNAs immobilized on filters were for 17 hr at 37°C in a solution containing  $10\%$  formamide,  $5\times$  standard saline phosphate EDTA (23),  $5 \times$  Denhardt's reagent (23), 0.1% (wt/vol) SDS,  $100 \mu$ g of sonicated and denatured salmon sperm DNA per ml, and  $2 \times 10^6$  cpm of probe oligonucleotide per ml. Posthybridization washing was for  $3 \times 30$  min in 800 ml of  $6 \times$ standard saline citrate  $(23)/0.1\%$  SDS at 37°C. Sixty-nine finger clones, in phage Agtl0, were arrayed on agar plates, and DNA was transferred to nylon filters and processed as in ref. 23. These were then screened with zfecl2 and washed at the appropriate stringency.

Isolation of Cloned DNAs. Insert cDNA was isolated from AgtlO by using <sup>a</sup> Cetus Amplitaq PCR kit and reaction conditions recommended by the manufacturers. The primers used were complementary to the DNA adjacent to the EcoRI cloning site in the vector. Primer sequences were ATGAG-TATTTCTTCCAGGGT and CAAGTTCAGCCTGGT-TAAGT. Cosmid DNAs were isolated by standard methods (23) from <sup>a</sup> library of 129/Sv mouse DNA (a gift of A.-M. Frischauf, London) and analyzed as described below. cDNA fragments for sequencing were cloned directly from the phage into the Bluescript plasmid vector (Stratagene) and sequenced by standard procedures (23). DNA sequence was obtained from both strands in all cases.

DNA Hybridizations. Mouse DNAs from the recombinant inbred (RI) strains were purchased from the Mouse DNA Resource of The Jackson Laboratory. Homozygous  $t^{w/8}$ 

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: RFLP, restriction fragment length polymorphism; cM, centimorgan(s); RI, recombinant inbred; SDP, strain distribution pattern.

<sup>\*</sup>To whom reprint requests should be addressed.

The sequences reported in this paper have been deposited in the GenBank data base [accession nos. M74235 (zfec12), M74236 (clone 18), M74237 (zfas8), and M74452 (zfec29)].

DNA was <sup>a</sup> gift of Gail Martin (San Francisco). DNA from the  $T/tf$  cross was a gift of Bernhard Herrmann (Tübingen). Southern analysis was by standard methods (23).

DNA for hybridization probes was isolated from zfecl2, zfas8, and zfec29 by isolation of the cDNA fragment from plasmid subclones by agarose gel electrophoresis. Probes from clone 18 were made directly from gel-purified PCRamplified material described above. DNAs were labeled by random priming to a specific activity of  $1-3 \times 10^9$  cpm/ $\mu$ g (24). Hybridization conditions were as in ref. 25 but omitting the bovine serum albumin. Posthybridization washes of filters were in  $1\%$  SDS/20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 6.5. Initially, filters were washed three times for 5 min in 800 ml of the buffer, which was prewarmed to 45 $\degree$ C, 55 $\degree$ C, or 65 $\degree$ C as appropriate, but the washes were not contained in a water bath. This was followed by washing for 20 min in a fresh prewarmed sample of the same buffer in a vessel contained in a shaking water bath at  $65^{\circ}$ C (high stringency),  $55^{\circ}$ C (moderate stringency), or  $45^{\circ}$ C (low stringency).

## **RESULTS**

Isolation of a Zinc Finger cDNA Encoded on Chromosome 17. We isolated <sup>a</sup> zinc finger cDNA clone, zfecl2, whose mRNA levels changed with the growth phase of <sup>a</sup> mouse lymphoblastoid cell line, L1210. zfecl2 was identified by oligonucleotide hybridizations (see Materials and Methods) and differential cDNA hybridization of isolated zinc fingercontaining cDNA clones (unpublished data). Partial sequence analysis identified at least 18 contiguous fingers at the C terminus of the predicted polypeptide. Northern analysis of L1210 mRNA show that zfecl2 cDNA detects an mRNA of  $\approx$ 3 kilobases (kb) (data not presented).

We identified a restriction fragment length polymorphism (RFLP) with Sst <sup>I</sup> between the DBA/2J and C57BL/6J inbred mouse strains and this enabled us to map zfec12 in the  $B \times D$ RI series to chromosome <sup>17</sup> (Fig. 1). We see linkage within 1.0 centimorgan (cM) (95% confidence limits of 0.00-4.53 cM; no recombinants in 24 RI strains) to the chromosome 17 loci, D17Lehl80, D17LehS4M, D17Leh443, D17Leh94 (21), and D17Leh111 and D17Leh550 (26). Bucan et al. (21) have shown that the first 4 loci are deleted in the  $t^{w18}$  partial t haplotype.

Identification of Related cDNAs and Genes. Southern blots ofwild-type mouse DNA suggested that there were additional genes related to zfecl2 in the genome. To isolate cDNAs for these genes, we screened 69 finger protein cDNAs by hybridization with zfecl2. At high stringency we see only



FIG. 1. B×D strain distribution patterns (SDPs) for zfec12, zfec29, and closely linked loci on chromosome 17. SDPs for zfecl2 and zfec29 were determined for RFLPs defined with Sst I. B denotes the C57BL/6J allele and D denotes the DBA/2J allele: for zfecl2, B  $= 25$  kb and  $D = 13$  kb; for zfec29,  $B = 3.1$  kb and  $D = 9.5$  kb. SDPs for DJ7Leh54M, DJ7Leh180, DJ7Leh443 and D17Leh94, and T66D and Hba-ps4 are taken from ref. 21; those for D17Leh111 and D17Leh550 are from ref. 26. An X indicates a crossover; a dash (-) indicates not determined.

homologous hybridization of zfecl2 to itself, but at reduced stringency we detect cross-hybridization of additional cDNAs: zfec29 and zfas8 cross-hybridize strongly, and zfec23 and zfec4 cross-hybridize more weakly (data not presented). Partial sequence analysis of zfec29 and zfas8 identifies multiple finger motifs, probably at the C-terminal end of the predicted polypeptides (data not presented).

Northern blot analysis of L1210 mRNA has not yet given us convincing data on the size of mRNAs that are homologous to zfec29 or zfas8. Neither probe detects abundant transcripts and we cannot exclude the possibility that the probes are cross-reacting with the more abundant zfecl2 mRNA.

Using an Sst <sup>I</sup> RFLP, zfec29 and zfecl2 have identical strain distribution patterns in the  $B \times D$  RI series (see Fig. 1). zfec4 was mapped to chromosome 7 and zfec23 was mapped to chromosome <sup>4</sup> (unpublished data). We were unable to identify a RFLP for zfas8.

Using the Sst <sup>I</sup> polymorphisms of zfecl2 and zfec29, we analyzed 58 recombinants between the  $T$  and  $tf$  region of chromosome 17 (B. Herrmann and H. Lehrach, personal communication). Both genes have identical segregation patterns and map 0.25 cM proximal to D17Leh54M and D17Lehl80 (the 443 and 94 markers have not yet been typed in this cross).

We screened <sup>a</sup> cDNA library constructed using mRNA from 8.5-day postcoitum embryos (27) with zfecl2, zfec29, and zfas8. As well as identifying further isolates of all three, we isolated an additional cDNA, clone 18, that was related but not identical. We conclude that the four cDNA clones represent members of a family of finger protein genes expressed at this stage of embryonic development (Fig. 2a-d).

In Fig. 3 we have compared DNA sequence from each cDNA clone that includes the codons representing the C-terminal 39 amino acids of the predicted translation product and the <sup>3</sup>' untranslated regions. Extensive sequence conservation exists between all four cDNAs. The pattern of sequence homology is not consistent with the cDNAs being derived from a single gene that has multiple alternative splices at the <sup>3</sup>' end.

We have isolated four apparently overlapping cosmid clones containing zfecl2, corresponding to 50 kb of DNA, and also related clones including an additional 80 kb. All of these cosmids hybridize strongly with the finger oligonucleotide (not presented). We also isolated cosmids containing



FIG. 2. Sequence relatedness of cDNAs and cosmids mapped to the  $t^{w18}$  deletion region: cDNA inserts from zfec12 (lane 1), zfec29 (lane 2), zfas8 (lane 3), and clone 18 (lane 4). (a) EtBr-stained gel.  $(b)$ Hybridization with the finger oligonucleotide.  $(c \text{ and } d)$  Hybridization with zfec12 at low stringency  $(c)$  and at high stringency  $(d)$ .  $(e)$ EcoRI-digested cosmid <sup>443</sup> DNA hybridized with the probe D17Leh443 (lane 5), the finger oligonucleotide (lane 6), and zfecl2 at low stringency (lane 7). Sizes are indicated in kb.

Genetics: Crossley and Little



FIG. 3. A 3' end DNA sequence comparison of four zinc finger protein cDNAs (A) and their C-terminal translation products (B). (A) Full sequence for zfec12. For the other cDNAs this is given only when it differs from that of zfec12. A period (.) marks sequences that are identical to zfec12; a dash (-) marks spaces inserted to maximize the alignment; bold type indicates coding DNA sequences; A<sup>n</sup> marks the start of the  $poly(A)$  tails. (B) An asterisk (\*) indicates conservative amino acid substitutions.

D17Leh180, D17Leh54M, and D17Leh443. Hybridization of these cosmids with the finger oligonucleotide shows strong



hybridization to the cosmid  $443$  (Fig. 2e) and weak hybridization to some of the fragments from cosmids 180 and 54M (not presented). zfec12 cross-hybridizes at reduced stringency with finger oligonucleotide hybridizing fragments contained in the 443 cosmid. Attempts to isolate zfec29, zfas8, and D17Leh94 cosmids have been unsuccessful.

Four Zinc Finger Genes Are Deleted in  $t^{w18}$ . Since D17Leh180, D17Leh54M, D17Leh443, and D17Leh94 are deleted in the mouse partial t haplotype  $t^{w18}$ , we wished to establish if the finger genes were also deleted. Martin et al. (28) derived a  $t^{w18}$  homozygous ES cell line, and in Fig. 4 we show that zfec12, zfec29, zfas8, and clone 18 are all deleted in the DNA of this isolate. We have also used a *t*-specific Sst I polymorphism of zfec12 to show that this gene is deleted (data not presented) in the partial  $t$  haplotype  $t^{ksl}$ , which is allelic with  $t^{w18}$  (21).

Similar Southern analyses of the homozygous  $t^{w18}$  ES cell line DNA, probed with zfec12 but washed at low stringency, are presented in Fig. 5, which shows that there are many zfec12-related fragments that are not deleted in  $t^{w18}$  homozygous DNA. We do not know if any of these related sequences map to chromosome 17 but outside of the  $t^{w18}$  deletion. Some of these fragments may be due to cross-hybridization to the chromosome 4 and 7 loci identified by the related clones zfec23 and zfec4.

## **DISCUSSION**

In summary, we have isolated four cDNA clones containing finger protein motif sequences and about 250 kb of genomic DNA. All sequences are derived from probes that have been shown to be deleted in the partial *t* haplotype  $t^{w18}$  and to be linked within 1.0-0.25 cM.

FIG. 4. At least four zinc finger protein genes are deleted in  $t^{w/8}$ :<br>wild-type C57BL/6J (B) and  $t^{w/8}/t^{w/8}$  (t) DNAs digested with EcoRI and hybridized with zfec12 (z12), zfec29 (z29), zfas8 (z8), and clone 18 (c18). Duplicate blots were washed at high stringency. All four sequences are lost from the  $t^{w/8}$  sample. Sizes of hybridizing fragments are in kb.



FIG. 5. Sequences related to zfecl2 are also present outside of the  $t^{w10}$  deletion. Wild-type  $(+)$  and  $t^{w10}/t^{w10}$  (t) EcoRI-digested DNAs were hybridized with zfecl2 and washed at moderate (a) or low stringency  $(b)$ . Hybridization is seen in the  $t^{w10}$  tracks, showing some sequences related to zfec12 are not deleted in  $t^{w18}$ . For comparison with Fig. 4, 4.4 marks the size in kb of the zfecl2 fragment.

Two pieces of evidence lead us to believe that the cDNAs are derived from separate genes and are not produced by alternative splicing of a single transcription unit. (i) Several cosmids contain separate regions that hybridize (or crosshybridize) with 3' or 5' specific probes derived from zfec12, implying that a complete transcription unit lies on a single cosmid. (ii) Our sequence analysis has failed to find identical regions shared between any of the cDNAs. Neither argument rigourously disproves alternative splicing, and definitive sequence analysis and mapping of the whole genomic region are necessary.

We believe that these genes are clustered at this site since the calculated density of random finger protein genes would be <sup>1</sup> every 4-10 Mb (300-700 genes), corresponding to an intergenic recombination frequency of 2-5 cM, assuming <sup>1</sup> cM is on average 2000 kb. Our linkage data are not compatible with random distribution. The genetic size of the  $t^{w18}$  deletion cannot be accurately determined. Analysis of the detailed map of this chromosomal region (26) makes some predictions: neither D17Leh111 nor D17Leh550 is deleted in  $t^{w18}$ , even though both map to the same  $B \times D$  recombinant interval as zfecl2 and zfec29 and the other deleted markers (ref. 21; unpublished data). The position of D17Leh111 is established in the Mus spretus interspecific cross (26) and shows D17Leh111 is proximal to D17Leh54M and must therefore be the marker closest to, and proximal of, the proximal breakpoint of the  $t^{w18}$  deletion.

In the M. spretus interspecific cross, D17Leh550 is linked to D17Leh180 and distal to D17Leh111. We have failed to identify RFLPs that would enable us to map D17Leh550 in any of the more informative crosses discussed here and we can conclude only that  $D17Leh550$  is either distal or proximal to the deletion. If it is distal, then the deletion size must be  $\leq$ 1.0 cM. If it is proximal, it is in the region that is between D17Leh111 and the proximal breakpoint. In this case, the nearest distal DNA marker known to be present in  $t^{w18}$  is Hba-4ps. However, the  $Fu$  and  $tf$  loci, which map proximal of Hba-4ps, are not revealed by  $t^{w18}$ . if maps  $\approx 2.5$  cM distal to D17Leh54M and D17Leh180 in the  $T/tf$  cross mentioned above. However,  $fu$  maps 2 cM proximal to  $tf$  and the distal

 $t^{w18}$  breakpoint must be in the small interval between the deletion markers and  $fu$ .

Preliminary analysis of yeast artificial chromosome (YAC) clones shows that all of the markers deleted in  $t^{w18}$ , except for D17Leh180, are contained on three overlapping clones and that all of the cross-hybridizing fragments detected by zfecl2 and known to be deleted in  $t^{w18}$  are contained on a single 470-kb fragment. D17Leh94 is proximal of the finger protein genes and DJ7LehS4M is distal. Based upon the assumption that all of the YACs are faithful representations of the genomic regions, we can conclude that the minimal size of the  $t^{w18}$  deletion must be 900 kb (data to be presented elsewhere).

The relationship of genes in the finger protein cluster and the lethal effect(s) is open to conjecture since the cluster must represent no more than half of the deleted DNA. The demonstrated properties of finger proteins make them of obvious interest for analysis of temporal and spatial distribution of their mRNAs, since this might give clues to their role, if any, in the complex pattern of disfunction that is characteristic of the  $tcl-4$  lethal. The loss of this region in the  $t^{w18}$  chromosome opens up immediate possibilities for complementation studies using transgenic animals: the size of the deletion will make this a challenging endeavor.

We are indebted to Bernhard Herrmann for his gift of the  $T/t$  cross DNAs, Gail Martin for her gifts of  $t^{w18}/t^{w18}$  DNA, Suzanne Gokool and Thomas Leung for isolation of clones, Keith Willison for  $t$ haplotype DNAs and advice, and Denise Barlow, Maja Bucan, and Alan Ashworth for helpful discussions. This work was supported by a grant from the Medical Research Council, United Kingdom.

- 1. Rhodes, D. & Klug, A. (1988) in Nucleic Acids and Molecular Biology, eds. Eckstein, F. & Lilly, D. M. J. (Springer, Heidelberg), pp. 149-165.
- 2. Gibson, T. J., Postma, J. P. M., Brown, R. S. & Argos, P. (1988) Protein Eng. 2, 209-218.
- 3. Ruiz <sup>i</sup> Altaba, A., Perry-O'Keefe, H. & Melton, D. A. (1987) EMBO J. 6, 3065-3070.
- 4. Miller, J., McLachlan, A. D. & Klug, A. (1985) EMBO J. 4, 1604-1614.
- 5. Brown, R. S., Sander, C. & Argos, P. (1985) FEBS Lett. 186, 271-274.
- 6. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987) Cell 51, 1079-1090.
- 7. Hartshorne, T. A., Blumberg, H. & Young, E. T. (1986) Nature (London) 320, 283-287.
- 8. Stillman, D. J., Bankier, A. T., Seddon, A., Groenhout, E. G. & Nasmyth, K. A. (1988) EMBO J. 7, 485-494.
- 9. Rosenberg, U. B., Schroeder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I. & Jaeckle, H. (1986) Nature (London) 319, 336-339.
- 10. Boulay, J. L., Dennefeld, C. & Alberga, A. (1987) Nature (London) 330, 395-398.
- 11. Tautz, D., Lehmann, R., Schnuerch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K. & Jaekle, H. (1987) Nature (London) 327, 383-389.
- 12. Moses, K., Ellis, M. C. & Rubin, G. M. (1989) Nature (London) 340, 531-536.
- 13. Morishita, K., Parker, D. S., Mucenski, M. L., Jenkins, N. N., Copeland, N. G. & Ihle, J. N. (1988) Cell 54, 831-840.
- 14. Kinzler, K. W., Ruppert, J. M. & Vogelstein, R. (1988) Nature (London) 332, 371-374.
- 15. Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A. & Spierer, P. (1990) Nature (London) 344, 219-223.
- 16. Joho, K. E., Darby, M. K., Crawford, E. T. & Brown, D. (1990) Cell 61, 293-300.
- 17. Bellefroid, E. J., Lecoq, P. J., Benhida, D. A., Poncelet, D. A. & Martial, J. A. (1989) DNA 8, 377-387.
- 18. Silver, L. M. (1985) Annu. Rev. Genet. 19, 179-208.
- 19. Hammer, M. F., Schimenti, J. & Silver, L. M. (1989) Proc. NatI. Acad. Sci. USA 86, 3261-3265.

Genetics: Crossley and Little

- 20. Klein, J., Sipos, E. & Figueroa, F. (1984) Genet. Res. 44, 39-46.
- 21. Bucan, M., Herrmann, B. G., Frischauf, A.-M., Bautch, V. L., Bode, V., Silver, L. M., Martin, G. R. & Lehrach, H. (1987) Genes Dev. 1, 376-385.
- 22. Snow, M. H. L. & Bennett, D. J. (1978) Embryol. Exp. Morphol. 47, 39-52.
- 23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1990) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
- 24. Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
- 25. Church, G. M. & Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991-1995.
- 26. Nadeau, J. H., Herrman, B. G., Bucan, M., Burkart, D., Crosby, J. L., Erhart, M. A., Kosowsky, M., Kraus, J. P., Michiels, F., Schnattinger, A., Tchetgen, M.-B., Varnum, D., Willison, K., Lehrach, H. & Barlow, D. (1991) Genomics 9, 78-89.
- 27. Fahner, K., Hogan, B. L. M. & Flavell, R. A. (1987) EMBO J. 6, 1265-1271.
- 28. Martin, G. R., Silver, L. M., Fox, H. S. & Joyner, A. R. (1987) Dev. Biol. 121, 20-28.