MOLECULAR AND CELLULAR BIOLOGY, Feb. 1987, p. 806–812
0270-7306/87/020806-07$02.00/0

Vol. 7, No. 2

Multiple Domains for the Chicken Cellular Sequences Homologous to the v-*ets* Oncogene of the E26 Retrovirus

ANNE GEGONNE,[1] DOMINIQUE LEPRINCE,[1] MARTINE DUTERQUE-COQUILLAUD,[2] BERNARD VANDENBUNDER,[1] ANNE FLOURENS,[1] JACQUES GHYSDAEL,[1] BRIGITTE DEBUIRE,[2] AND DOMINIQUE STEHELIN[1]*

Laboratoire d'Oncologie Moléculaire, Unité 186 Institut National de la Santé et de la Recherche Médicale-Centre National de la Recherche Scientifique 1160, Institut Pasteur, 59019 Lille Cédex,[1] and Unité 124 Institut National de la Santé et de la Recherche Médicale, Institut de Recherches sur le Cancer, 59045 Lille,[2] France

Received 11 August 1986/Accepted 28 October 1986

We have investigated the structure of chicken genomic DNA homologous to v-*ets*, the second cell-derived oncogene of avian retrovirus E26. We isolated a c-*ets* locus spanning ca. 30.0 kilobase pairs (kbp) in the chicken genome with homologies to 1,202 nucleotides (nt) of v-*ets* (total length, 1,508 nt) distributed in six clusters along 18.0 kbp of the cloned DNA. The 5′-distal part of v-*ets* (224 nt) was homologous to chicken cellular sequences contained upstream within a single 16.0-kbp *Eco*RI fragment as two typical exons but not found transcribed into the major 7.5-kb c-*ets* (or 4.0-kb c-*myb*) RNA species. Between these two v-*ets*-related cellular sequences we found ca 40.0 kbp of v-*ets*-unrelated DNA. Finally, the most 3′ region of homology to v-*ets* in the cloned DNA was shown to consist of a truncated exon lacking the nucleotides coding for the 16 carboxy-terminal amino acids of the viral protein but colinear to one of the two human c-*ets* loci, c-*ets-2*.

Avian retrovirus E26 induces a mixed erythroid-myeloid leukemia predominantly of the erythroid type (13). In addition to v-*myb*, the genome of E26 contains a second stretch of cell-derived specific sequences, v-*ets* (1,508 nucleotides [nt]) 3, 16, 19). The cellular counterpart of v-*ets*, c-*ets*, is clearly distinct from c-*myb* since it is transcribed in some normal cell types as a major 7.5-kilobase (kb) polyadenylated RNA (as well as two minor 2.2- and 1.5 (2.0)-kp species) in contrast to the 4.0-kb c-*myb* mRNA (16, 27). Furthermore, the human cellular genes c-*myb* and c-*ets* have been assigned to different chromosomes: 6q22-24 for c-*myb* (14), 11q23-24 for c-*ets-1* (8, 26), and 21q22.1-22.3 for c-*ets-2* (26). The E26 viral transforming protein is a 135,000-M_r fusion protein, P135$^{gag\text{-}myb\text{-}ets}$ translated from the 5.7-kb genomic viral RNA and located in the nucleus of E26-transformed cells (2, 4, 5, 15). v-*ets* is not related to any other known viral oncogenes but its deduced amino acid sequence displays homology with the products of yeast cell cycle genes *CDC4* and *CDC36* (20).

In this report we demonstrate that the v-*ets* sequences originated mainly from two stretches of DNA 40.0 kbp distant. We isolated lambda recombinant phages corresponding to a c-*ets* locus transcribed as a major 7.5-kb mRNA and containing six exons homologous to 1,202 nt of a central portion of v-*ets*. In addition, the 5′ part of v-*ets* was acquired from two exons located in a 16.0-kbp *Eco*RI fragment of chicken DNA which lies within the c-*ets* locus described here. These two exons do not appear to be transcribed into the major 7.5-kb c-*ets* RNA species seen in several cell types. Finally, by sequencing analysis of the most 3′ cellular sequence homologous to v-*ets*, we identified a putative translation stop codon to this c-*ets* locus, colinear with that of the human c-*ets-2* locus but distinct from the one used for the termination of the E26 viral protein. Thus, v-*ets* appears to represent a mosaic of probably at least three stretches of DNA.

* Corresponding author.

MATERIALS AND METHODS

Preparation of DNAs. High-molecular-weight chicken DNA was prepared from total embryos as previously described (23). Plasmid DNA and recombinant phage DNAs were prepared as described previously (22).

Hybridization probes. Lambda E26Q1 is a molecularly cloned E26 provirus. A 2.5-kbp *Eco*RI fragment containing the v-*ets* domain flanked 5′ and 3′ by E26 v-*myb* and *env* sequences, respectively, was subcloned in plasmid pKH47 to yield the plasmid clone pVE2.5 (Fig. 1) (16). The different v-*ets* and v-*myb* probes used in this study were prepared from the pVE2.5 clone.

Fragments representative of the viral *env* gene were derived from the Prague A strain of Rous sarcoma virus (Pr-RSVA) genome as illustrated by Saule et al. (22). These fragments were labeled by nick-translation with New England Nuclear Corp. nick translation kits.

Isolation of recombinant phages containing c-*ets* chicken sequences. A library of chicken erythrocyte DNA fragments in Charon 4A vectors was kindly provided by J. Dodgson and J. Engel (9). A total of 2.5×10^5 plaques were screened with a v-*ets* probe by the in situ plaque purification procedure of Benton and Davis (1).

A recombinant DNA library in vector EMBL4 constructed with a partial *Sau*3A digest of total chicken embryo DNA by A. Begue in our lab was used to isolate lambda clones containing sequences homologous to the 5′ part of v-*ets*.

Restriction mapping and gel electrophoresis of poly(A)-containing cellular RNAs. Restriction endonucleases were obtained from Boehringer Mannheim and Bethesda Research Laboratories. Digested DNA fragments were size separated by electrophoresis in horizontal agarose gels and transferred to nitrocellulose filters by the method of Southern (24). Total cellular RNA was fractionated on oligo(dT)-cellulose (T3; Collaborative Research). Polyadenylated RNAs were denatured, size separated, and transferred to nitrocellulose by the method of Thomas (25).

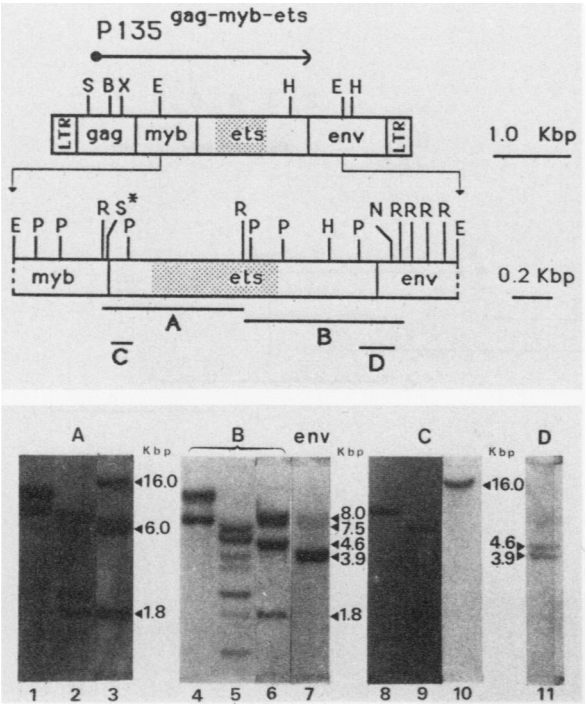


FIG. 1. (Top) Schematic drawing of the E26 provirus deduced from restriction endonuclease mapping and partial sequencing analysis of recombinant phages isolated from DNA libraries of two distinct E26-infected cells (22, 26, 28). Abbreviations: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; S, *Sst*I; X, *Xho*I. *gag*, *env*, and LTR (long terminal repeat) refer to helper virus-related sequences (13), whereby *myb* and *ets* refer to the two cell-derived sequences present in the E26 genome (16, 19). The arrow indicates the 3.6-kbp of coding capacity required for the synthesis of the fusion protein $P135^{gag\text{-}myb\text{-}ets}$ (4, 5, 15). (Middle) Restriction endonuclease map of the 2.5-kbp *Eco*RI fragment subcloned into plasmid pKH47 (pVE2.5). This fragment contains all the v-*ets* sequences flanked 5′ and 3′ by E26 v-*myb* and *env* sequences, respectively. Abbreviations: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; N, *Nde*I; R, *Rsa*I. S* represents a convenient *Sau*3A restriction site located at the E26 *myb-ets* junction. The stippled area represents the regions of homology found between v-*ets* and the yeast cell cycle genes *CDC4* and *CDC36* (20). (Bottom) The different v-*ets* probes (A, B, C, and D) derived from the pVE2.5 subclone are indicated. After digestion of chicken chromosomal DNA by *Bam*HI (lanes 1, 4, and 8), *Hin*dIII (lanes 2, 5, and 9), or *Eco*RI (lanes 3, 6, 7, 10, and 11) the DNA fragments were separated by electrophoresis in the same 0.8% agarose gel and transferred onto nitrocellulose sheets (24). These Southern blots were hybridized with the v-*ets* probes indicated above: probe A (lanes 1 to 3), probe B (lanes 4 to 6), probe C (lanes 8 to 10), and probe D (lane 11). Similarly, the chicken DNA was hybridized with an *env* probe (lane 7), since probes B and D were shown to contain such sequences (19).

Hybridization of blots to ^{32}P-labeled DNA, washing, and autoradiography at −70°C with Kodak X-ray films and Du Pont Lightning-Plus X-ray intensifying screens were performed as described previously (22).

Heteroduplex mapping. The lambda chicken c-*ets* DNAs were hybridized with either E26 recombinant phages containing the v-*ets* sequence in the same orientation versus the lambda vector arms: lambda E26-1, which has been described (16), or lambda E26-2 containing the 9.0-kbp *Bam*HI fragment from lambda E26-1 but cloned in the opposite orientation into the lambda 47 vector. The heteroduplex molecules spread onto the hypophase were picked up on the parlodion-coated grids (300 mesh) and stained with uranyl acetate. The grids were subsequently dried, rotary shadowed, and viewed in a Hitachi HU12 electron microscope.

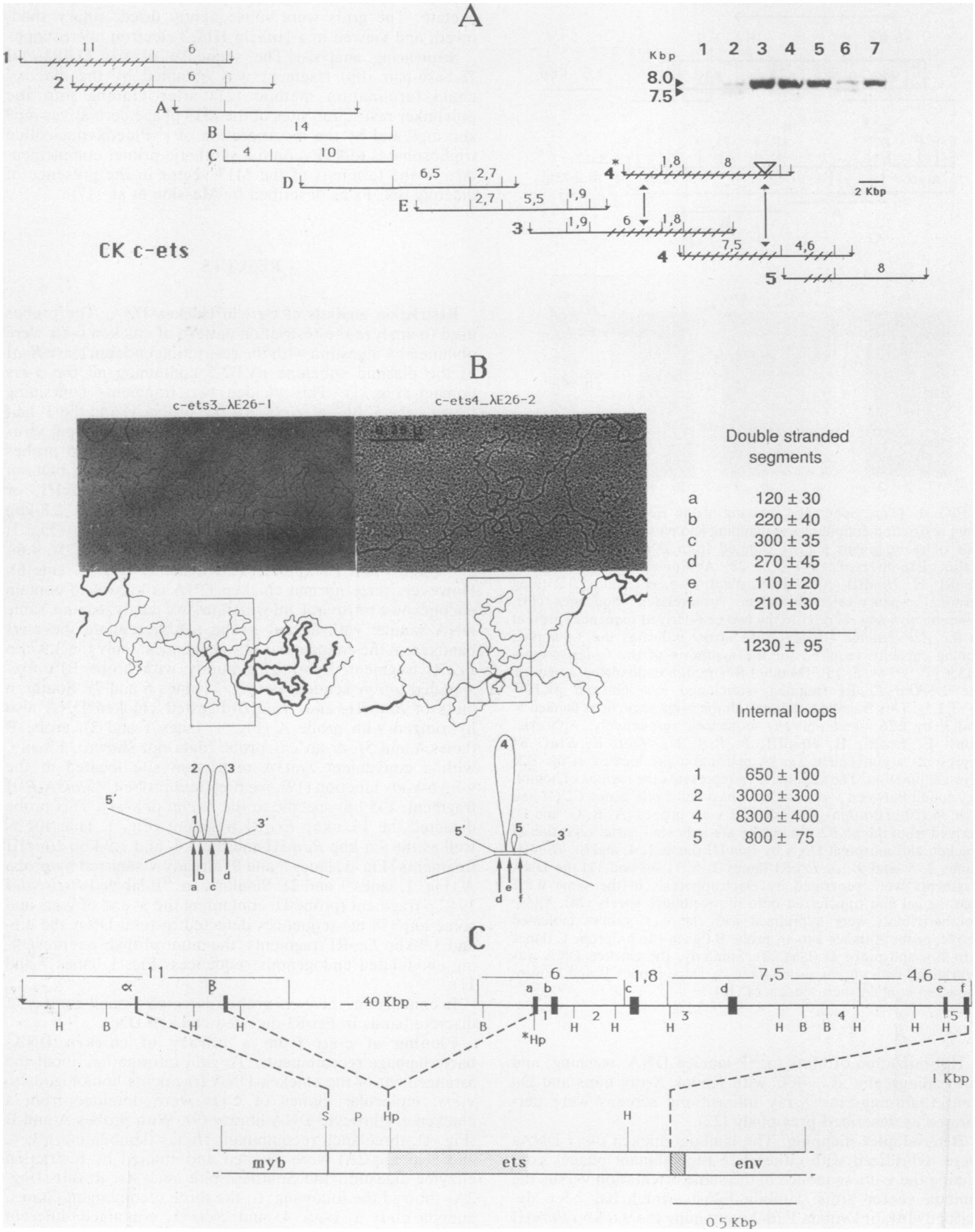
Sequencing analysis. The sequence of the *Eco*RI-*Pst*I 78-base-pair (bp) fragment was obtained by the dideoxy chain termination method (21) after cloning into the polylinker restriction sites of the M13 phage derivatives mp8 and mp9 and by the incorporation of [^{32}P]deoxynucleotide triphosphates (dNTPs) onto a synthetic primer complementary to the terminus of the M13 vector in the presence of dideoxy dNTPs as described by Messing et al. (17).

RESULTS

Restriction analysis of c-*ets* in chicken DNA. The probes used to analyze the restriction pattern of chicken c-*ets* were obtained by digestion with the restriction endonuclease *Rsa*I of the plasmid subclone pVE2.5 containing all the v-*ets* domain (Fig. 1). This yielded two fragments containing roughly the 5′ half of v-*ets* (768 bp, probe A) and the 3′ half of v-*ets* (765 bp) together with 105 bp related to helper virus *env* sequences (870 bp, probe B) (19). These two probes hybridized to several DNA fragments in Southern blots of chicken DNA digested with either *Eco*RI, *Bam*HI, or *Hin*dIII. After *Eco*RI digestion, 16.0-, 6.0- and 1.8-kbp *Eco*RI fragments hybridized with the 5′ v-*ets* probe (Fig. 1, lane 3), whereas the 3′ v-*ets* probe detected 8.0-, 7.5-, 4.6-, 3.9- (faint), and 1.8-kbp *Eco*RI fragments (Fig. 1, lane 6). However, since normal chicken DNA is known to contain endogenous retroviral information, we analyzed the same DNA sample with an *env* probe to differentiate the c-*ets* bands from the endogenous *env* sequences. Only the 3.9-kbp *Eco*RI fragment (hybridizing faintly with probe B) corresponded to *env* sequences (Fig. 1, lanes 6 and 7). Southern blots of *Bam*HI- and *Hin*dIII-digested chicken DNA also hybridized with probe A (Fig. 1, lanes 1 and 2), probe B (lanes 4 and 5), or an *env* probe (data not shown). Finally, with a convenient *Sau*3A restriction site located at the v-*myb*-v-*ets* junction (19), we prepared probe C (*Sau*3A-*Pst*I fragment, 135 bp) specific to the 5′ end of v-*ets*. This probe detected the 16.0-kbp *Eco*RI fragment (Fig. 1, lane 10) as well as the 7.5-kbp *Bam*HI and the 5.4- and 1.9-kbp *Hin*dIII fragments (Fig. 1, lanes 8 and 9) already visualized by probe A (Fig. 1, lanes 1 and 2). Similarly, a ^{32}P-labeled *Pst*I-*Nde*I 194-bp fragment (probe D) containing the 3′ end of v-*ets* and some *env* (78 nt) sequences detected in total DNA the 4.6- and 3.9-kbp *Eco*RI fragments, the latter probably representing *env*-related endogenous sequences (Fig. 1, lanes 7 and 11).

In conclusion, the v-*ets* probes detected a small number of discrete bands in *Eco*RI-digested chicken DNA.

Cloning of c-*ets* from a library of chicken DNA-bacteriophage recombinants. To gain information about the arrangement of the chicken DNA fragments homologous to v-*ets*, molecular clones of c-*ets* were obtained from a chicken erythrocyte DNA library (9). With probes A and B (Fig. 1), three such recombinant clones (lambda c-*ets* 3, 4, and 5 in Fig 2A) were isolated and studied by restriction enzyme digestion and Southern blot analysis. Results (Fig. 2A) showed the following: (i) the three recombinant clones, lambda c-*ets* 3, c-*ets* 4, and c-*ets* 5, contained different overlapping cellular inserts; (ii) none of these recombinant phages hybridized with the 5′-most v-*ets* probe (probe C) or


A
CK c-ets
Kbp 1 2 3 4 5 6 7
8.0
7.5
2 Kbp
B
c-ets3_λE26-1
c-ets4_λE26-2
Double stranded segments
a 120 ± 30
b 220 ± 40
c 300 ± 35
d 270 ± 45
e 110 ± 20
f 210 ± 30
1230 ± 95
Internal loops
1 650 ± 100
2 3000 ± 300
4 8300 ± 400
5 800 ± 75
C
40 Kbp
*Hp
1 Kbp
myb
ets
env
0.5 Kbp

contained part of the 16.0-kbp *Eco*RI fragment detected in cellular DNA by this probe (Fig. 1, lane 10); and (iii) lambda c-*ets* 4 and c-*ets* 5 hybridized with the 3′ v-*ets* probe (probe D in Fig. 1) and contained the 4.6-kbp *Eco*RI fragment (Fig. 1, lane 11). These 3′ v-*ets*-homologous sequences were flanked downstream by 8.0 kbp of v-*ets*-unrelated chicken DNA in lambda c-*ets* 5. Moreover, the regions of homology to v-*ets* in lambda c-*ets* 3 and c-*ets* 4 DNAs were precisely localized by heteroduplex analysis; six small double-stranded regions separated by five internal loops could be defined (Fig. 2B).

Incidentally, we also isolated from the same DNA library a recombinant phage named lambda c-*ets* 4* which seemed to overlap with both lambda c-*ets* 3 and lambda c-*ets* 4 since each ^{32}P-labeled *Eco*RI fragment of lambda c-*ets* 4* cross-hybridized with a corresponding fragment in one of the two clones (Fig. 2A). However, lambda c-*ets* 4* and c-*ets* 4 contained two *Eco*RI fragments different in size (7.5 and 8.0 kbp) which were both detected by a v-*ets* probe in the chicken DNA used in the experiments shown in Fig. 1. This is probably due to allelic variation in this region of c-*ets*, since analysis of DNA obtained from seven individual chicken embryos revealed either both *Eco*RI fragments or only the 8.0-kbp *Eco*RI fragment (Fig. 2A).

In conclusion, cellular sequences corresponding to the major part of v-*ets* (1,202 nt; see below) are split into six regions of homology (presumably exons), interrupted by five regions of nonhomology (introns) within ca. 18.0 kbp of chicken DNA (Fig. 2) cloned in lambda c-*ets* 3, 4, and 5. It should be noted that both the *ets-1* and *ets-2* sequences found split on different chromosomes in human, cat, and mouse DNA (26) can be found contiguous in chicken cellular DNA.

Cloning of the cellular sequences homologous to the 5′ part of v-*ets*. From a newly constructed library of *Sau*3A-digested chicken embryo DNA, we isolated two recombinant phages (named lambda c-*ets* 1 and lambda c-*ets* 2) hybridizing with the 5′-specific v-*ets* probe C. Lambda c-*ets* 1, although containing a single 11.0-kbp artificial *Eco*RI fragment instead of the 16.0-kbp *Eco*RI fragment (Fig. 1, lane 10), contained the 7.5-kbp *Bam*HI as well as the 5.4- and 1.9-kbp *Hin*dIII fragments detected in total chicken DNA (Fig. 1, lanes 8 and 9, and Fig. 2C). These *Hin*dIII fragments contained cellular sequences, named α and β, accounting for the 5′ part of v-*ets*. Nucleotide sequence analysis revealed that α and β exhibited the typical properties of eucaryotic exons, with open reading frames (82 and 142 nt, respectively) identical to that used by E26 virus and flanked by consensus splice signals (Fig. 3) (18). Nevertheless, we found no splice acceptor sequence at the leftward boundary of α or upstream sequence homology to v-*myb*. Thus, α represents an exon probably truncated during an illegitimate recombination with *myb* sequences. Appropriate splicing between α and β would yield the most 5′ sequences of v-*ets*. To determine more precisely the relationship between the two genomic regions homologous to v-*ets*, we also sequenced exon a, the first v-*ets*-related exon of lambda c-*ets* 3 (Fig. 2). From the results shown in Fig. 3, it appears that the most obvious mechanism to generate the v-*ets* sequences involves a consensus splicing of β to a.

These results as well as the fact that lambda c-*ets* 1 and lambda c-*ets* 3 did not contain overlapping DNA inserts raised the question whether they represented two distinct regions of the chicken genome or distant parts of a same locus. We thus extensively screened chicken DNA libraries with ^{32}P-labeled fragments of lambda c-*ets* 1 and lambda c-*ets* 3. Using in addition the "walking gene" method, we isolated five recombinant clones (lambda c-*ets* A, B, C, D, and E) containing v-*ets*-unrelated inserts totaling ca. 40 kbp which represented the junction between the two v-*ets* homologous cellular stretches (Fig. 2). The transcription of c-*ets* in chicken cells as a major 7.5-kb RNA species (16, 27) most likely correlates with a large cellular gene. E26 could therefore have transduced and joined in its genome two sequences, namely those found in lambda c-*ets* 1 and c-*ets* 3, that may belong to a same locus but are separated by a large region (40.0 kbp) of cellular DNA. To test this hypothesis, polyadenylated RNAs of MSB1 cells (a chicken line of T lymphoblasts transformed by Marek disease virus, expressing a high level of 7.5-kb c-*ets* RNA [16]) was hybridized with probe C. No signal was observed (Fig. 4, lane C1). The lack of hybridization was significant, since the same blot hybridized with the ^{32}P-labeled 870-bp *Rsa*I fragment (probe B) revealed the typical c-*ets* RNA species (Fig. 4, lane B1); furthermore, probe C readily detected the 5.7-kb E26 genomic RNA as efficiently as did probe B (Fig. 4, lanes 2). To bolster this demonstration we analyzed MSB1 mRNAs with probe E (*Pst*I-*Pst*I. fragment, 420 bp) which encompasses *myb* sequences and probe C; only the 4.0-kb c-*myb* mRNA species was detected (Fig. 4, lane E1). In addition, when we used probe A, which represents the 5′ end of v-*ets* (including α and β), only the c-*ets* mRNA species described above with probe B were detected (Fig. 4, lane A1). Therefore, the 5′ part of v-*ets* is homologous to chicken cellular sequences which are not transcribed into the major 7.5-kb c-*ets* RNA species in MSB1 cells. However, we cannot rule out the possibility that the 5′ part of v-*ets* may be expressed by an alternate splicing mechanism as part of a minor c-*ets*

FIG. 2. Structure of the c-*ets* locus. (A) Molecular clones of c-*ets* were obtained from different chicken (CK) DNA libraries: *Alu*I-*Hae*III-digested erythrocyte DNA, *Eco*RI-digested fibroblast DNA, and *Sau*3A-digested total embryonic DNA (see the text). Vertical bars represent true *Eco*RI fragments, and arrows represent artificial *Eco*RI fragments due to the construction of some of the libraries used. Fragments containing sequences homologous to v-*ets* (diagonal lines) in clones lambda c-*ets* 1, 2, 3, 4, 4*, and 5 were identified by Southern blot analysis with v-*ets* probes (data not shown). Recombinant clones A, B, C, D, and E which did not hybridize with v-*ets* probes were isolated by the walking gene technique. Orientation is 5′ to 3′. Chromosomal chicken DNAs obtained from different embryos were restricted with *Eco*RI. The DNA fragments were separated by agarose gel electrophoresis, transferred to nitrocellulose sheets, and hybridized with the ^{32}P-labeled 8.0-kbp *Eco*RI fragment of lambda c-*ets* 4*. (B) Heteroduplex analysis of chicken c-*ets* recombinant clones. Top to bottom: Representative heteroduplex, interpretative sketch (thick lines represent double strands and thin lines represent single strands), and schematic drawing of the contour lengths (features 1 to 5 refer to single strands [loop 3 was 3200 ± 300] and features a to f refer to double strands, whose lengths in base pairs are summarized in the table). (C) Restriction map of a c-*ets* chicken locus. The sites of cleavage of the following restriction endonucleases into lambda c-*ets* 1, 3, and 4 are shown: *Bam*HI (B), *Eco*RI (E), and *Hin*dIII (H). In addition the position of one *Hpa*I site (Hp*) is shown as it is also found in the viral v-*ets* sequence. The 5′ and 3′ limits of the homology found between v-*ets* and the second c-*ets* domain have been roughly determined by heteroduplex analyses. These studies revealed six regions of homology (a to f) separated by five regions of nonhomology (1 to 5) (see the table). The hatched box represents the most 3′ sequences of v-*ets* which are not derived from the c-*ets* exon f and remain of unknown origin. The stippled box represents the homology with *CDC4* and *CDC36* (20).

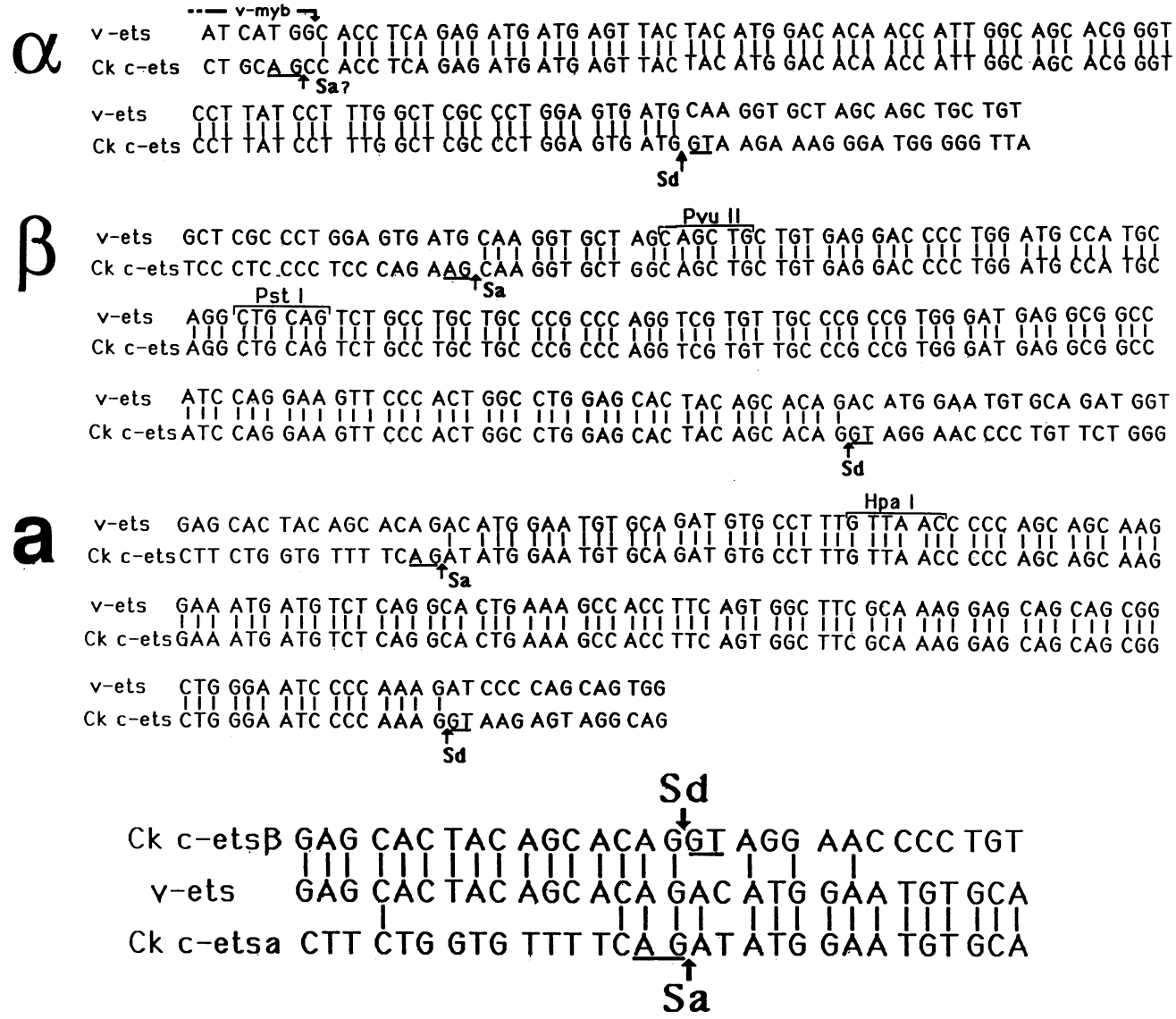


FIG. 3. Nucleotide sequence of α, β, and a. Nucleotide sequence of relevant portions from lambda c-*ets* 1 and c-*ets* 3 clones is aligned with the published v-*ets* sequence (19). Three open reading frames (82, 142, and 120 nt) colinear to v-*ets* (with a single base change at the beginning of exon a) and flanked by typical splice signals (18) with the exception of the leftward boundary of α (see the text) were detected. The v-*ets* sequence probably results from splicing of α and β to a. Ck, Chicken; Sd, splice donor; Sa, splice acceptor.

RNA species present at a very low level in this cell line or expressed in so far unidentified cell types.

v-*ets* and c-*ets* differ at their 3′ end. Sequencing analysis of v-*ets* and the topography of the E26 provirus suggested that the same translation stop codon was used for termination of synthesis of the viral protein P135$^{gag\text{-}myb\text{-}ets}$ and the cellular c-*ets* protein P54$^{c\text{-}ets}$ (6, 11, 19). Therefore, we investigated in more detail the cellular sequences homologous to the 3′ end of v-*ets*. From hybridization experiments with total chicken DNA and with chicken DNA recombinant clones (Fig. 1, lanes 6, 7 and 11, and Fig. 2), these sequences should be localized within the 4.6-kbp *Eco*RI fragment of clones lambda c-*ets* 4 or lambda c-*ets* 5. A *Pst*I-*Eco*RI 78-bp fragment hybridizing with the 3′-distal v-*ets* probe D (Fig. 1) was isolated from the 4.6-kbp *Eco*RI fragment of lambda c-*ets* 4, subcloned into phage M13mp8, and sequenced by the chain termination method (21). This sequence was compared with the previously published v-*ets* sequence (19). A perfect homology beween the v-*ets* and the c-*ets* fragment was found (Fig. 5), but only for the first 34 nt following the *Pst*I site. No significant homology was found with the last 82 nt of v-*ets*. However, the chicken c-*ets* sequence described here is colinear, with some minor base changes probably due to phylogenetic divergence, with that of a human c-*ets* cDNA clone corresponding to human c-*ets-2* (Fig. 5) (27). Therefore, the viral v-*ets* sequence on one hand and the sequences of human c-*ets-2* cDNA and chicken c-*ets* DNA on the other hand appear to specify different carboxy termini for their encoded proteins. It should also be noted that from sequence analysis (Fig. 5), probe D, although containing 78 *env*-related nt, 34 nt derived from the 4.6-kbp *Eco*RI fragment of lambda c-*ets* 4 but also 82 nt of unknown origin (see sequence data in Fig. 5), hybridized only with two cellular *Eco*RI fragments (4.6 and 3.9 kbp in size) with similar

intensities (Fig. 1, lane 11). A possible explanation could be that the 4.6-kbp band (lane 6) represented a doublet of a first fragment located within the lambda c-*ets* 4 recombinant clone and derived from the c-*ets* locus and a second fragment which could contain the termination codon used by E26 for its transforming protein. The nature and origin of this latter fragment remain open questions.

DISCUSSION

In this paper, we investigate the nature and structure of the cellular sequences homologous to the E26-specific cell-derived domain v-*ets* (1,508 nt). Surprisingly, the cellular chicken sequences accounting for the v-*ets* domain appeared to be split into two stretches (6.0 kbp and 18.0 kbp, respectively) separated by about 40.0 kbp of v-*ets*-unrelated DNA. The downstream stretch related to v-*ets* contained six clusters of homology, named a to f (Fig. 2), scattered along 18.0 kbp of DNA. Therefore, the *ets-1* and *ets-2* loci which appeared to be located on different chromosomes in several mammalian species, including humans, are found contiguous in chickens, the species from which the E26 virus arose. However, at least two sets of *ets*-related proteins can be observed in chicken tissues: $P54^{c\text{-}ets}$, which is highly related to the v-*ets* domain of $P135^{gag\text{-}myb\text{-}ets}$ (6, 11), and the P60, P62, and P64 proteins, which have only limited homology with $P54^{c\text{-}cts}$ and $P135^{gag\text{-}myb\text{-}cts}$ (12). We do not know whether these two sets of *ets*-related proteins are translated from differentially spliced mRNA transcripts derived from the locus described here or from two distinct loci, one being highly related to v-*ets*, the other displaying only a small region of homology. The characterization of the nucleotide sequence of the corresponding cDNAs should resolve the issue.

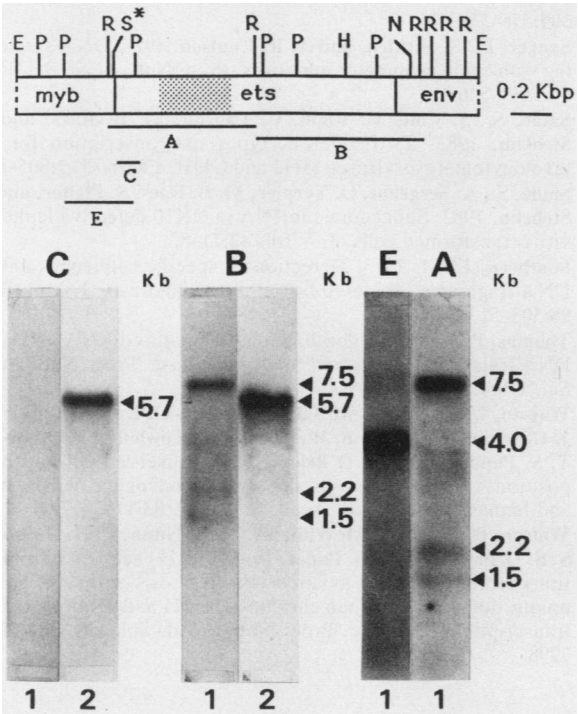


FIG. 4. The 5′ part of v-*ets* was not found transcribed into the c-*ets* or c-*myb* RNAs. Polyadenylated RNAs of MSB1 (a T-lymphoid cell line) (lanes 1) or of E26-transformed myeloblasts (lanes 2) were denatured with glyoxal-dimethyl sulfoxide, separated by agarose gel electrophoresis, bound to nitrocellulose sheets (25), and hybridized with the indicated probes (C, B, E, and A). See the Legend to Fig. 1 for symbols and abbreviations.

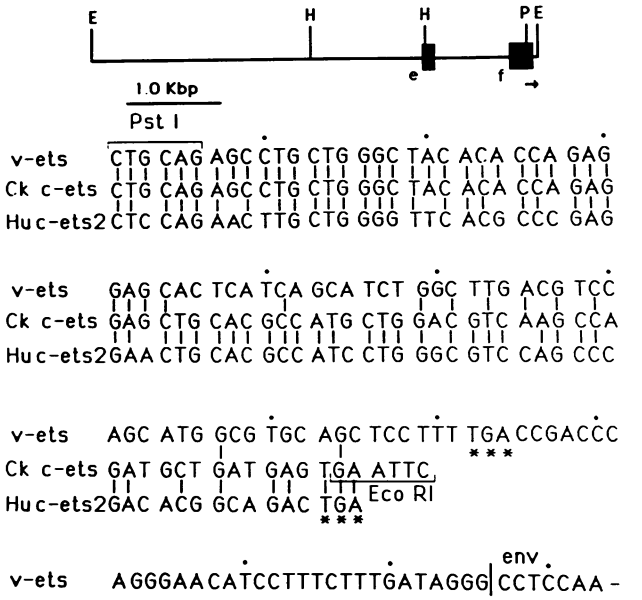


FIG. 5. P135 translational stop codon not derived from c-*ets* exon f. A 78-nt c-*ets* fragment that includes the 3′ stretch of homology between v-*ets* and the 4.6-kbp *Eco*RI fragment lambda c-*ets* 4 recombinant clone (top) was subjected to nucleotide sequence analysis. Numbering starts at the C residue of a *Pst*I restriction site (nt 1927 in reference 27). Sequencing was done as described by Sanger et al. (21) after cloning of a *Pst*I-*Eco*RI fragment into phage M13 derivatives mp8 and mp9. ***, Translation stop codon used for synthesis of the E26 transforming protein. The border with the *env* gene is also shown. The human c-*ets-2* cDNA sequence with its translation stop codon is also indicated (27). Vertical bars refer to homologous nucleotides found within the three sequences. Ck, Chicken; Hu, human.

The 5′ c-*ets* stretch contains sequences homologous to the first 224 nt of v-*ets*. This 5′ c-*ets* stretch, although displaying a split structure typical of a eucaryotic gene (Fig. 3), is not transcribed into the 7.5-kb major RNA species (Fig. 4). This suggests that these sequences could be expressed in minor RNA species present at low levels (just above the limit of sensitivity of the method we used) in the MSB1 cellular RNAs tested or in so far unidentified cell types. However, antisera directed against these sequences corresponding to the amino-terminal v-*ets*-encoded domain, although identifying the P135 viral protein, failed to precipitate any of the described cellular c-*ets* proteins in many different cell types (10). Finally, in support of these results, the beginning of homology found between v-*ets* and the yeast cell cycle genes *CDC4* and *CDC36* (20) seems to correlate with the upstream part of the second cellular stretch of homology to v-*ets* defined in this study.

Turning now to the 3′ end of v-*ets*, the topography of E26 proviral DNA as well as sequencing analysis strongly suggested that the translation stop codon used for the synthesis of the viral P135 protein was of cellular origin and derived from c-*ets* (16, 19). However, sequencing analysis of the cellular sequence homologous to the most 3′ v-*ets* sequence in the 4.6-kbp *Eco*RI fragment of lambda c-*ets* 4 demonstrated that (i) the last 16 amino acids of the $P135^{gag\text{-}myb\text{-}ets}$ protein are not derived from this part of c-*ets* and (ii) the region of homology (named f in Fig. 2C) found between c-*ets* and v-*ets* displays features typical of a truncated c-*ets* exon

and is colinear (with some single base changes) to a human c-*ets-2* cDNA which contains a translation stop codon. Therefore, this f probably provides the termination codon of at least one of the chicken c-*ets* proteins, namely P54 (11).

Using specific oligonucleotide probes, we are currently cloning the cellular sequences containing the translation stop codon of the E26 transforming protein and isolating c-*ets* cDNAs. These studies will shed light on the mechanism by which three probably distinct cellular domains have been transduced in E26 virus and fused to encode a single viral protein.

ACKNOWLEDGMENTS

We thank J. Coll and S. Saule for stimulating discussions and critical reading of this manuscript; A. Fritsch, G. Torpier, and A. Jacob for help with the heteroduplex analysis; J. Dogson for the chicken erythrocyte DNA library; A. Begue for the *Sau*3A chicken embryo DNA library; C. Lagrou for expert technical assistance in cell culture; and N. Devassine for patient typing.

This work was supported by the Institut National de la Santé et de la Recherche Médicale (U186-U-124), Centre National de la Recherche Scientifique (UA 04 1160), Association pour la Recherche sur le Cancer, and the Pasteur Institute of Lille.

LITERATURE CITED

1. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones to hybridization to single plaques in situ. Science **196**:180–182.
2. **Beug, H., M. S. Hayman, and T. Graf.** 1982. Myeloblasts transformed by the avian acute leukemia virus E26 are hormone-dependent for growth and for expression of a putative *myb*-containing protein, P135 E26. EMBO J. **1**:1069–1073.
3. **Bishop, J. M.** 1983. Cellular oncogenes and retrovirus. Annu. Rev. Biochem. **52**:301–334.
4. **Bister, K., M. Nunn, C. Moscovici, B. Perbal, M. A. Baluda, and P. Duesberg.** 1982. Acute leukemia viruses E26 and avian myeloblastosis virus have related transformation specific RNA sequences but different genetic structures, gene products and oncogenic properties. Proc. Natl. Acad. Sci. USA **79**:3677–3681.
5. **Boyle, W. S., M. A. Lampert, J. S. Lipsick, and M. A. Baluda.** 1984. Avian myeloblastosis virus and E26 virus oncogene products are nuclear proteins. Proc. Natl. Acad. Sci. USA **81**:4265–4269.
6. **Chen, J. H.** 1986. The proto-oncogene c-*ets* is preferentially expressed in lymphoid cells. Mol. Cell. Biol. **5**:2993–3000.
7. **Coll, J., M. Righi, C. de Taisne, C. Dissous, A. Gegonne, and D. Stehelin.** 1983. Molecular cloning of the avian acute transforming retrovirus MH2 reveals a novel cell-derived sequence (*v-mil*) in addition to the *myc* oncogene. EMBO J. **2**:2189–2194.
8. **de Taisne, C., A. Gegonne, D. Stehelin, A. Bernheim, and R. Berger.** 1984. Chromosomal localization of the human proto-oncogene *c-ets*. Nature (London) **310**:581–583.
9. **Dodgson, J. B., J. Strommer, and D. Engel.** 1979. Isolation of the chicken B globin gene and a linked embryonic B-like globin gene from a chicken DNA recombinant library. Cell **17**:879–887.
10. **Gegonne, A., P. Pognonec, D. Leprince, D. Dernis, E. Remaut, D. Stehelin, and J. Ghysdael.** 1986. Preparation and characterization of specific antisera directed against different polypeptidic domains encoded by the *v-ets* oncogene of the avian acute leukemia virus E26. C. R. Acad. Sci. (Paris) **303**:253–256.
11. **Ghysdael, J., A. Gegonne, P. Pognonec, D. Dernis, D. Leprince, and D. Stehelin.** 1986. Identification and preferential expression in thymic and bursal lymphocytes of a *c-ets* protooncogene encoded Mr 54 000 cytoplasmic protein. Proc. Natl. Acad. Sci. USA **83**:1714–1718.
12. **Ghysdael, J., A. Gegonne, P. Pognonec, K. Boulukos, D. Leprince, D. Dernis, C. Lagrou, and D. Stehelin.** 1986. Identification in chicken macrophages of a set of proteins related to, but distinct from, the chicken cellular *c-ets* encoded protein $P54^{c\text{-}ets}$. EMBO J. **5**:2251–2256.
13. **Graf, T., and D. Stehelin.** 1982. Avian leukaemia viruses oncogenes and genome structure. Biochem. Biophys. Acta **651**:245–271.
14. **Harper, M. E., G. Franchini, J. Love, M. I. Simon, R. C. Gallo, and F. Wong-Staal.** 1983. Chromosomal sublocalization of human *c-myb* and *c-fes* cellular onc genes. Nature (London) **304**:169–171.
15. **Klempnauer, K. H., G. Symonds, G. I. Evan, and J. M. Bishop.** 1984. Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken c-*myb* gene. Cell **37**:537–547.
16. **Leprince, D., A. Gegonne, J. Coll, C. de Taisne, A. Schneeberger, C. Lagrou, and D. Stehelin.** 1983. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. Nature (London) **306**:395–397.
17. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
18. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
19. **Nunn, M. F., P. H. Seeburg, C. Moscovici, and P. H. Duesberg.** 1983. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. Nature (London) **306**:391–395.
20. **Peterson, T. A., J. Yochem, B. Byers, M. F. Nunn, P. H. Duesberg, R. F. Doolittle, and S. I. Reed.** 1984. A relationship between the yeast cell cycle genes CDC4 and CDC36 and the *ets* sequence of oncogenic virus E26. Nature (London) **309**:556–558.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminator inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Saule, S., J. Coll, M. Righi, C. Lagrou, M. B. Raes, and D. Stehelin.** 1983. Two different types of transcription for the myelocytomatosis viruses MH2 and CMII. EMBO J. **2**:805–809.
23. **Saule, S., A. Sergeant, G. Torpier, M. B. Raes, S. Pfeifer, and D. Stehelin.** 1982. Subgenomic mRNA in OK10 defective leukemia virus transformed cells. J. Virol. **42**:71–82.
24. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
25. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
26. **Watson, D. K., M. S. McWilliams-Smith, C. Kozak, R. Reeves, J. Gearhart, M. F. Nunn, W. Nash, J. R. Fowle III, P. Duesberg, T. S. Papas, and S. J. O'Brien.** 1986. Conserved chromosomal positions of dual domains of the *ets* protooncogene in cats, mice and humans. Proc. Natl. Acad. Sci. USA **83**:1792–1794.
27. **Watson, D. K., M. J. McWilliams, M. F. Nunn, P. H. Duesberg, S. S. O'Brien, and T. S. Papas.** 1985. The *ets* sequence from the transforming gene of avian erythroblastosis virus, E26, has unique domains on human chromosomes 11 and 21: both loci are transcriptionally active. Proc. Natl. Acad. Sci. USA **82**:7294–7298.