Mammalian genomes contain numerous fragments of DNA that are derived from inactivated transposable elements. The accumulation and persistence of these elements is generally attributed to transposase activity rather than through possession or acquisition of a function of value to the host genome. Here we describe such a repetitive element, named ALF (for annexin VI LINE-2 fragment), comprising 130 bp of DNA derived from a LINE-2 sequence, which functions as a potent T-cell-specific silencer. The expansion of the DNA database arising as a result of the human genome sequencing project enabled us to identify ALF in, or close to, several well characterized genes including those for annexin VI, interleukin-4 and protein kinase C-β. A systematic analysis of the entire LINE-2 sequence revealed that ALF, and not other regions of the LINE-2 sequence, was especially highly represented in the human genome. Acquisition of a function by this repetitive element may explain its abundance. These data show that a conserved fragment of an interspersed nuclear element has the potential to modulate gene expression, a discovery that has broad implications for the way in which we view so-called 'junk' DNA and our understanding of eukaryotic gene regulation.

INTRODUCTION

Eukaryotic gene transcription is largely under the control of sequence-specific DNA-binding proteins that may act proximally or distally to transcription start sites. Most proteins of this type have been reported to enhance or increase the rate of transcription, but a significant number of transcriptional inhibitors or repressors have now also been described (1). Typical gene promoters contain numerous binding sites for such positive and negative regulatory factors but, because the binding sites are short, it is unusual to find extended regions of sequence similarity between different promoters. The best characterized conserved non-coding sequences are Alu sequences and long interspersed nuclear elements (LINEs) retrotransposons (2), both of which occur in abundance in mammalian genomes.

In humans, Alu and mammalian-wide interspersed repeat (MIR) sequences together with LINEs comprise >25% of the entire genome (3). Of these three types of repetitive element, LINEs are more closely associated with having a functional role. A small proportion of LINEs are active, full-length and have been reported to be capable of retrotransposition (4). Recently, an extended MIR sequence was described known as LINE-2 (3). In this study, we describe a LINE-2 fragment, which commonly is located proximally to or within the introns of many human genes. We have used the sequence from the annexin VI (ANX6) promoter as a model to investigate the function of this element and show that it is a potent T-cell-specific silencer. These findings are consistent with the idea that persistence of this element in mammalian evolution may be because it has a function of value to the host genome, rather than because of any intrinsic retrotransposase activity.

RESULTS AND DISCUSSION

We recently completed a functional analysis of the human ANX6 gene promoter (5). Extending this work, we found that the full-length promoter was almost an order of magnitude weaker than the minimal promoter in Jurkat T cells. This is in contrast to epithelial and fibroblast lines, in which the full-length ANX6 promoter has similar activity to the minimal active promoter. This suggested that an element(s) located distally within the ANX6 promoter might be repressive in T cells, but not in other cell types. To find out whether sequences in the ANX6 promoter were related to sequences known to be involved in transcriptional silencing, we used overlapping 200 bp fragments of the ANX6 promoter to search the non-redundant nucleotide database at the NCBI using BLAST. Unexpectedly, a region of the promoter extending from bases –712 to –586 were similar to a large number (>100) of sequences, all of which were identified using RepeatMasker (http://ftp.genome.washington.edu/RM/RepeatMasker.html ) as LINE-2 fragments. Six of these sequences were aligned with the ANX6 sequence using MegAlign, and a consensus was derived in which inclusive nucleotides are defined as those present in more than half of the aligned sequences (Fig. 1a). The human and bovine interleukin-4 (IL-4) sequences both conform to the consensus, but the former was excluded to make the point that this element is not restricted to the human genome. Note that the ANX6 sequence is 5' to 3' whereas the IL-4 sequence is reverse complement. Table 1 gives more information concerning the origins and loci of these elements, many of which are within introns of known or hypothetical genes.

To find out whether other similar sized LINE-2 fragments were represented at this level in the human genome, we used
overlapping 150 bp regions extending over the entire LINE-2 sequence to search for homologous sequences (Fig. 1b). Most of these 150 bp sequences were poorly represented, but the region corresponding to the domain in the ANX6 promoter occurred with high frequency. The 3' end of the LINE-2 sequence is an MIR that is already known to be an abundant genomic element (3). The ANX6 LINE-2 fragment (named ALF) was found within or close to several well characterized genes, including IL-4 and protein kinase C-β (PKC-β) (Fig. 1c). Interestingly, the ANX6 and IL-4 genes are neighbours within the cytokine gene cluster on chromosome 5q (6,7). This may be coincidence, but the existence of conserved regulatory elements within a cluster of genes may facilitate coordinated regulation of gene expression during cell differentiation or activation.

The conservation of ALF might be due either to preferential transposition as an element in its own right, or to preferential retention over surrounding LINE-2 sequence because it has a property of some value to the host genome. Fortuitously, orthologous ALFs were identified in the human and bovine IL-4 genes so it was possible to test whether these ALFs, derived from a common ancestral ALF, had diverged more or less than...
neighbouring non-coding sequences. If surrounding non-coding sequence shows similar levels of homology to ALF this argues against preferential retention. Figure 1d shows a diagonal plot of these orthologous regions and reveals that ALF does in fact stand out (together with regions x and y) as a conserved sequence in this region of the IL-4 intron 3. Thus ALF, x and y have been conserved in preference to surrounding sequences. Further comparison of the human and bovine IL-4 ALFs with the LINE-2 consensus (Fig. 1e and f) shows that ALF and an extended region of surrounding sequences, including x and y, are all of LINE-2 origin, suggesting that they represent the remains of a complete ancestral LINE-2 element. It is interesting to note that regions x and y also correspond to more highly represented parts of the LINE-2 sequence as shown in Figure 1b. ALF thus did not enter the IL-4 gene as a repeating element in its own right, but as part of a LINE-2 element. Therefore, a LINE-2 element in the ancestral IL-4 gene has, during evolution to the human and bovine genes, preferentially retained several conserved regions, of which ALF is one.

We then investigated the second possible reason for the abundance of ALF, namely that ALF has a function that may be useful to the host genome. In IL-4 and PKC-β, ALF was located some distance from the respective promoter and in reverse and forward orientation, respectively (Fig. 1c). To test whether ALF had genuine silencer activity, we cloned the sequence into the pGL3-promoter, in forward and reverse orientation, distally and proximally to the promoter (Fig. 2). In both orientations and irrespective of distance from the promoter, ALF virtually abolished the production of luciferase in Jurkat cells, thereby satisfying the criteria for definition as a silencer (1,8). However, in A431 and HeLa cells, ALF functioned as both enhancer and silencer, depending on orientation, position and cell type. In HeLa cells, ALF was weakly active in the forward orientation when proximal to the promoter, but had strong enhancer activity (3- to 8-fold higher) in the reverse orientation, whether distal or proximal to the promoter. In A431 cells, ALF functioned as a weak enhancer when proximal to the promoter, but repressed transcriptional activity when located distally to the promoter.

To gain further insight into the repressive activity observed in T cells, we first examined a short but highly conserved 20 nucleotide domain (the HR1 domain in Fig. 1a) at the 5' end of ALF. Deletion of this region, generating the construct ΔHR1-Luc, substantially relieved the repressive activity observed in Jurkat cells, particularly when cloned in the forward orientation (Fig. 3a). In A431 and HeLa cells, removal of HR1 led to loss of enhancer activity. These data indicate that HR1 contributes to the dual role of ALF. However, the fact that ΔHR1-Luc was still repressive in Jurkat cells indicates that HR1 is not solely responsible for the silencer activity of ALF. The mechanisms by which silencers suppress promoter activity are not well understood, but they may act by directly interfering with the basal transcription complex (9). This type of interaction was suggested to underlie transcriptional repression exerted by the silencer in the 3'-untranslated region of the IL-4 gene in Th1 cells (10). Interestingly, although there was no evidence of T-cell-specific transcription factors binding to the HR1 domain, we did find myc-associated zinc finger (MAZ) binding (11) to HR1 (data not shown) and, because MAZ binding has been linked to DNA bending (12), it is possible that ALF could function via a direct interaction with the basal transcription complex. However, the MAZ-binding site is not present in

Table 1. Chromosomal location and characteristics of genes containing ALF

<table>
<thead>
<tr>
<th>Sequence</th>
<th>GenBank accession no.</th>
<th>Region of homology</th>
<th>% homology to αxVI region</th>
<th>Source</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AC002398</td>
<td>33 237–33 108</td>
<td>50.8</td>
<td>Human Chr 19q13.1</td>
<td>Intron 6 of α-chimaerin homologue</td>
</tr>
<tr>
<td>2</td>
<td>AG000382</td>
<td>71–203</td>
<td>50.0</td>
<td>Human Chr 21q</td>
<td>No further information available</td>
</tr>
<tr>
<td>3</td>
<td>U141159</td>
<td>889–764</td>
<td>52.4</td>
<td>Bos taurus</td>
<td>Intron 3 of the IL-4 gene</td>
</tr>
<tr>
<td>4</td>
<td>AC004030</td>
<td>7388–7246</td>
<td>54.8</td>
<td>Human Chr 19p13.3</td>
<td>Intron 3 of an unknown hypothetical protein</td>
</tr>
<tr>
<td>5</td>
<td>AC002299</td>
<td>83 094–82 961</td>
<td>46.8</td>
<td>Human Chr 16p12</td>
<td>Intron 5 of the PKC-β gene</td>
</tr>
<tr>
<td>6</td>
<td>Z95113</td>
<td>72 854–72 721</td>
<td>52.4</td>
<td>Human Chr 22q11.2</td>
<td>Nearest identified EST 12 kb upstream</td>
</tr>
</tbody>
</table>

Figure 2. ALF is a T-cell-specific silencer. ALF was cloned into pGL3-promoter in forward (F) and reverse (R) orientations, proximally and distally to the promoter. Constructs were tested for activity in Jurkat, HeLa and A431 cells. For each data set, the activity of pGL3-promoter was set at 100%. Data are the means of four (Jurkat) or six (HeLa and A431) independent transfections (±SEM) normalized by dot-blotting.
Because IL-4 gene expression in T cells can be modulated by phorbol ester and calcium ionophore (10,13–15), we performed electrophoretic mobility shift assays (EMSAs) with HR1 comparing nuclear extracts from control cells with those prepared from cells exposed to these agonists for 2 h. The results in Figure 3b show that HR1 forms a gel-shifted complex in activated Jurkat cells that is absent in A431, HeLa or Daudi cells. Note that HR2 did not form specific gel-shifted complexes in similar experiments, and HR3 and HR4 remain to be investigated. Figure 3c shows that addition of cycloheximide to the cells at the same time as the agonists inhibited induction of the HR1 complex, whereas cyclosporin A had no effect. These results show that the induced factor required de novo gene transcription and was not a resident cytosolic transcription factor such as a STAT (signal transducer and activator of transcription) or NF-AT (nuclear factor of activated T cells). To define the binding site for this factor more closely, we searched the Transfac database (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl) for known transcription factor-binding sites. The only recognized sequence in this short domain, CCCTCCC, corresponded to the MAZ-binding site (11). We then investigated binding to a series of DNA oligomers with unitary deletions in both the 5' and 3' directions and resolved the binding site to the 12 nucleotide sequence, 5'-CTCCCACTCACT-3'. Since this sequence does not correspond to the binding site for any known transcription factor, binding may occur via a novel protein induced by phorbol ester and ionophore. Finally, we investigated the effects of phorbol ester and calcium ionophore on expression of the endogenous ANX6 gene. The western blot in Figure 4a shows that the level of annexin VI protein falls rapidly between 8 and 12 h after addition of the agonists. Annexin I, blotted using the same extracts, maintained a constant level of expression during the course of the experiment. These findings were corroborated by northern blotting (Fig. 4b), which showed that annexin VI mRNA falls to barely detectable levels 12 h after the application of phorbol ester and calcium ionophore. Thus, like the IL-4 gene, which also contains an HR1 sequence within ALF, ANX6 gene expression is down-regulated by phorbol ester and calcium ionophore in T cells.

In summary, we have identified a LINE-2 fragment named ALF that is a potent T-cell-specific silencer. We also show that agonists that down-regulate ALF-containing genes in T cells induce a factor that binds to a sequence within ALF. These findings are in contrast to other reports associating enhancer or promoter activities with repetitive elements (16,17), because ALF has the potential to function as a cell-type-specific silencer. We favour the hypothesis that this is not an arbitrary activity, and that ALF contributes to gene regulation in vivo. This work thus supports the idea that fragments of a repetitive element may be preserved preferentially within the host genome during evolution if they take on a function that is advantageous to the host.

MATERIALS AND METHODS

Materials

Tissue culture media and supplements were from Life Technologies. Western-Blue substrate, goat anti-rabbit alkaline phosphatase-conjugated antibody and the luciferase assay kit were
from Promega. [γ-32P]ATP (3000 Ci/mmol) and [α-32P]dCTP (3000 Ci/mmol) were from Dupont-NEN. DNA oligomers were from Life Technologies. Hybnd-N membrane was from Amersham and Immobilon-P membrane was from Millipore. All other chemicals were obtained from Sigma. pGL3-basic and pGL3-promoter luciferase vectors were from Promega. TA cloning vectors were from Promega. pTAG cloning vector was from R&D Systems. Hybond-N membrane was from Amersham. All other DNA-modifying enzymes were purchased from Promega.

Annexin VI promoter constructs

The deletion series of the ANX6 promoter in pGL3-basic has been described elsewhere (5). The 126 bp repetitive element was amplified by PCR using the oligonucleotides 5'-CTCCCTCCCTACTCTC-3' and 5'-TACCTGGGAAAAAGTGTAC-3' as upstream and downstream primers, respectively, with the ANX6 promoter as template (18). The first of these primers (annealed to its complement) was also used as the HR1 oligo in gel shift assays (see later). The PCR product was subcloned in forward and reverse orientations into pGL3-promoter, both proximally and distally (3 kb upstream) to the SV40 promoter. To test the function of the most 5' proximal region, the upstream primer was also used as the HR1 oligo in gel shift assays (see later). The PCR product was subcloned into pGL3-promoter in forward and reverse orientations proximally to the SV40 promoter. All constructs were verified by DNA sequencing.

Cell culture and transfections

A431 and HeLa cells were cultured and transfected as described previously (5). Jurkat and Daudi cells were cultured in RPMI containing 10% heat-inactivated fetal calf serum, 25 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin sulfate. Jurkat cells were transfected by electroporation using a GenePulser II electroporator (Bio-Rad). A total of 107 cells were harvested per transfection, pelleted at 1200 r.p.m. for 5 min at room temperature and washed once in electroporation buffer (0.2 M HEPES pH 7.4, 0.14 M NaCl, 0.15 M KCl, 7 × 10^-4 M NaHPO4, 6 × 10^-3 M glucose). Cells were then resuspended in 10 ml of electroporation buffer and pelleted again, before being resuspended in 250 µl of electroporation buffer. DNA (25 µg) was added and the cells were transferred to 4 mm electroporation cuvettes (Bio-Rad). Electroporation was performed at 400 V, 125 µF and 3000 V. Cells were allowed to recover for 5 min at room temperature and then transferred into 10 ml of complete growth medium. Luciferase activity was measured as described previously (5).

Dot blotting was performed as described previously (5). EMSAs were performed as described elsewhere (5,19) using control cells or cells that had been stimulated with 100 ng/ml phorbol 12-myristate 13-acetate and 1 µM A23187. Unlabelled double-stranded HR1 oligonucleotide (5'-CTCCCTCCCTACTCTC-3') and HR2 oligonucleotide (5'-TCCAGCCACACTGGCCCCCTGCTG-3') were used as competitor and non-competitor, respectively. Cycloheximide and cyclosporin A were used at concentrations of 10 and 500 ng/ml, respectively. Northern and western blotting were as described previously (5).

Abbreviations

ALF, annexin VI LINE-2 fragment; IL-4, interleukin-4; LINE, long interspersed nuclear element; PKC, protein kinase C.

Acknowledgements

This work was supported by the Arthritis Research Council, the Wellcome Trust, the Medical Research Council and GlaxoWellcome.

References