Genomic Analysis Reveals a Novel Nuclear Factor-κB (NF-κB)-binding Site in Alu-repetitive Elements

Received for publication, February 24, 2011; in revised form, August 6, 2011. Published, JBC Papers in Press, September 5, 2011, DOI 10.1074/jbc.M111.234161

Athina Antonaki‡1, Constantinos Demetriades§1, Alexander Polyzos‡1, Aggelos Banos‡1, Giannis Vatsellas‡, Matthieu D. Lavigne§, Effie Apostolou§, Eva Mantouvalou§, Deppie Papadopoulou§, George Mosialos§, and Dimitris Thanos‡

From the ‡Institute of Molecular Biology, Genetics and Biotechnology, Biomedical Research Foundation, Academy of Athens, 4 Soranou Efesiou Street, Athens 11527, Greece and the §Department of Biology, Aristotle University of Thessaloniki, University Campus, Thessaloniki 54124, Greece

Background: The transcription factor NF-κB is a critical regulator of immune response pathways.

Results: 11% of total NF-κB binding events in humans occur at the Alu-κB element present in Alu repeats.

Conclusion: NF-κB has a primate-specific function and a role in human evolution.

Significance: Repetitive elements expand the repertoire of binding sites to engage new genes into regulatory networks.

The transcription factor NF-κB is a critical regulator of immune responses. To determine how NF-κB builds transcriptional control networks, we need to obtain a topographic map of the factor bound to the genome and correlate it with global gene expression. We used a ChIP cloning technique and identified novel NF-κB target genes in response to virus infection. We discovered that most of the NF-κB-bound genomic sites deviate from the consensus and are located away from conventional promoter regions. Remarkably, we identified a novel abundant NF-κB-binding site residing in specialized Alu-repetitive elements having the potential for long range transcription regulation, thus suggesting that in addition to its known role, NF-κB has a primate-specific function and a role in human evolution. By combining these data with global gene expression profiling of virus-infected cells, we found that most of the sites bound by NF-κB in the human genome do not correlate with changes in gene expression of the nearby genes and they do not appear to function in the context of synthetic promoters. These results demonstrate that repetitive elements interspersed in the human genome function as common target sites for transcription factors and may play an important role in expanding the repertoire of binding sites to engage new genes into regulatory networks.

The transcription factor NF-κB orchestrates gene expression programs controlling the response to a variety of stimuli, including bacterial and viral infections and diverse biological functions, such as inflammation, cell proliferation, cellular homeostatic mechanisms, and carcinogenesis (1). In vertebrates, NF-κB is a mixture of homo- and heterodimers of a family of five sequence-specific DNA-binding protein subunits that include p50 (NF-κB1), p52 (NF-κB2), p65 (RelA), RelB, and c-Rel (2). Access of NF-κB to DNA elements is inhibited by a family of inhibitory molecules called IκBs (IκBα, IκBβ, IκBγ, IκBε, IκBζ, p105, and p100) (3, 4). Two major pathways of NF-κB activation have been characterized (5). The canonical pathway culminates with the nuclear translocation of the p50/p65 heterodimer, whereas the non-canonical pathway depends primarily on the activity of the p52/RelB heterodimer. The canonical pathway plays an important role in innate and adaptive immune responses, whereas the non-canonical pathway is critical for lymphoid tissue development, organization, and function (1, 2, 5, 6).

NF-κB proteins work as transcriptional activators or repressors by binding to the consensus DNA sequence 5’-GGGRNYYYCC-3’ known as the κB site. There are ~10^6 consensus κB sites in the human genome, and ~500 genes are regulated by NF-κB (a list of all known NF-κB target genes has been compiled by T. Gilmore; see the Boston University NF-κB Transcription Factors Web site) (reviewed in Refs. 7 and 8). However, the actual number of putative κB sites in the human genome is much greater because variations that deviate from the consensus motif are also bound by NF-κB and have been also shown to play critical roles in gene transcription despite the wide variation in NF-κB DNA binding affinity (9–15). Several studies have provided evidence that transcription factors, like NF-κB, can bind to many non-consensus sites with affinities similar to the ones displayed for consensus sites (reviewed in Ref. 16) and that most of the binding events in the nucleus occur in extragenic regions located far away from enhancers and promoters (17–21). However, these previous studies did not test for any significance of NF-κB binding and have raised a number of additional questions. For example, what proportion of the κB sites is functional? How do the κB sites that regulate gene expression compete with those that are not involved in gene regulation? Are there κB sites without functional importance?
Recently, we have established the biological significance of NF-κB interaction with three κB sites localized on specialized Alu-repetitive elements and showed that this binding is required for the subsequent delivery of NF-κB to target enhancers (e.g. IFN-β) via interchromosomal interactions (22). However, we do not know whether NF-κB binds selectively only to these three Alu repeats of thousands existing in the human genome; nor do we know the fraction of the activated NF-κB molecules bound to the Alu repeats. These early observations raised additional questions regarding the property of NF-κB and possibly of other transcription factors to bind repetitive DNA sequences under physiological conditions. Repetitive DNA represents a major component of non-coding sequences in higher eukaryotes and comprises ~60% of the human genome (23–26). Alu repeats are part of the short interspersed nuclear elements with an average size of 300 bp and exist in ~1,500,000 copies in the cell. Several previous studies have implicated short interspersed nuclear elements in regulation of gene expression (22, 24–26). How do transcription factors select specific sites for DNA binding among the many similar ones in the genome, especially the ones occurring on repetitive DNA elements? Do transcription factors, like NF-κB, bind preferentially to the higher affinity consensus sites compared with the larger number of lower affinity sites? What is the role of the chromatin state in making accessible transcription factor binding sites throughout the genome? Does binding of transcription factors in remote sites cause local chromatin remodeling and/or recruitment of the basal transcription machinery? Does activation of a transcription factor like NF-κB by different signals induce selective binding to some sites, or is site selection signal-independent? Are all occupied sites by NF-κB capable of supporting transcriptional activation? Put simply, is NF-κB binding restricted to only the sites that control gene expression? If not, what is the functional significance of the other sites? These are obvious questions given the relatively small number of NF-κB-regulated genes (~500) and the huge number of putative κB sites in the human genome (~2 × 10^9).

In this study, we have addressed many of the above questions using the NF-κB family of transcription factors as a model system to understand basic regulatory principles of transcription factor-dependent gene expression networks. It has been previously hypothesized that specificity in NF-κB-mediated transcriptional response can, in part, be provided through the combinatorial effects of the five different Rel members that are selectively regulated by κB proteins and activated by distinct signals in different cell types (reviewed in Ref. 7) and by the promoter context of the κB sites (27–29). Here, we identified a novel NF-κB-binding site residing in specialized Alu-repetitive elements that have the potential for long range transcription regulation, thus suggesting that in addition to its known role, NF-κB has a primate-specific function and a role in human evolution. We found that ~11% of the total NF-κB binding events in human cells occur at the Alu-κB site. By combining these data with global gene expression profiling of virus-infected cells, we found that most of the sites bound by NF-κB in the human genome, including the Alu-κB sites, do not obviously correlate with changes in gene expression of the nearby genes. These results demonstrate that repetitive elements interspersed in the human genome are common target sites for transcription factors and may play an important role in expanding the repertoire of binding sites to engage new genes into regulatory networks.

**EXPERIMENTAL PROCEDURES**

**Cell Cultures, Transfections, and Plasmid Construction—**
HeLa cells were maintained in DMEM containing 10% FBS, transfected using the calcium phosphate method, and infected with Sendai virus as described previously (29). IB4 cells are Epstein-Barr virus (EBV)-transformed human B cells, and they were maintained in RPMI 1640 medium containing 10% FBS. For the Western blot of Fig. 7B, HeLa cells were transfected on 6-well plates, harvested in SDS-PAGE sample buffer, and probed with the antibodies p65 (sc-109) and β-actin (sc-130301) from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA).

The thymidine kinase-luciferase constructs were generated by subcloning the genomic fragments from the p65 library. The plasmids (pCRII-TOPO vector) were digested with ScaI/XhoI and cloned upstream of the thymidine kinase promoter driving the expression of the luciferase reporter gene (30). The −40IFN-βluc constructs were generated by subcloning tandem repeats of oligonucleotides bearing the Alu-κB-binding site (one, two, or four copies) from the p65 library clones and placed upstream of the IFN-β core promoter driving the expression of the luciferase reporter gene (29). The DNA FISH experiments were carried out as described previously (22).

**Chromatin Immunoprecipitation and Library Construction—**

Chromatin preparation and immunoprecipitation experiments were carried out as described previously (31). The following antibodies were used and obtained from Santa Cruz Biotechnology, Inc.: p65 (sc-109), p52 (sc-298), RelB (sc-28689), PolII (sc-9001), and TBP (sc-273). The following histone-specific antibodies were obtained from Upstate Biotechnology: H4ac (06-866) and H3ac (06-599). The primers and oligonucleotides used in this study are indicated in supplemental Fig. 6.

To construct the p65 ChIP library, we immunoprecipitated chromatin prepared from HeLa cells infected with Sendai virus for 6 h, and the purified DNA fragments were processed with the tagged random PCR method, as described previously (32). Following transformation, we picked 450 clones and obtained sequence from 366 clones. **Supplemental Fig. 1** shows the sequence of each clone along with its genomic location and the sequence of the putative κB site. The sequence from the 366 clones were analyzed by BLAST to locate the genomic position and by TRANSFAC to identify putative NF-κB-binding sites.

The library containing the p52-binding sites was prepared by immunoprecipitated chromatin from IB4 cells with the p52 antibody, and the purified DNA fragments were treated with terminal deoxynucleotidyltransferase (New England Biolabs) to add poly(A) tails to the 3’ ends. The isolated DNA fragments were enriched by PCR using oligo(dT) primers containing BamHI flanking overhangs. This DNA mix was subsequently digested with BamHI and cloned into the BamHI site of pBlue-ScriptII vector.

**In Vitro Transcription and EMSAs—** Recombinant NF-κB (p50/p65 heterodimer) was expressed and purified as described.
A Novel NF-κB-binding Site in Alu-repetitive Elements

Previously (33). Similarly, EMSAs and in vitro transcription reactions were carried out as described previously (33).

**RESULTS**

**Virus Infection Induces a Wide Distribution of NF-κB DNA Binding into the Human Genome**—To investigate the localization of NF-κB into the human genome following virus infection, we applied a chromatin immunoprecipitation-based cloning technique to isolate genomic DNA sequences bound by NF-κB dimers using specific antibodies against the p65 (RELA) subunit. The ChIP-PCR-cloning procedure was optimized to avoid nonspecific interactions and bias in DNA cloning as verified by IgG control immunoprecipitation experiments and by applying limited PCR amplification to limit background (see “Experimental Procedures”). HeLa cells were mock- or virus-infected with Sendai virus for 6 h, followed by chromatin preparation and immunoprecipitation with the anti-p65 antibody (31). The precipitated DNA was processed as described previously (32). Four hundred fifty colonies were randomly picked, and the inserts were sequenced, producing 366 DNA sequences that were characterized in detail. The lengths of the cloned DNA fragments ranged from 100 to 850 bp, averaging about 430 bp (data not shown).

We searched the human genomic data base by BLAST to determine the location of the p65-bound DNA fragments in relation to known or predicted genes (supplemental Fig. 1). Fig. 1A shows the distribution of the identified NF-κB-binding sites relative to the closest transcription start site of nearby genes. Interestingly, the larger percentage of NF-κB-binding sites (49.5%) was found in regions located more than 10 kb away from known or predicted regulatory enhancer/promoter regions (Fig. 1A). Furthermore, ~34% of NF-κB-binding sites were identified in introns, whereas smaller percentages ranging from 1.5 to 9% corresponded to κB sites located in exons, 5′- and 3′-untranslated regions, and 5′-proximal regions (less than 10 kb upstream from known genes). Notably, in ~38% of the cloned NF-κB-bound fragments, a consensus or a putative NF-κB-binding site could not be identified (Fig. 1B). Overall, these findings are consistent with several previous reports in which transcription factor-bound DNA sites were mapped into mammalian genomes using ChIP-cloning, ChIP-PET, and ChIP-seq techniques (reviewed in Ref. 16). Our results are in agreement with previous studies reporting binding of NF-κB into the human genome in response to LPS activation of macrophages or monocytes (10, 14, 36) or treatment of HeLa cells with TNF-α and mapping of NF-κB binding across human chromosome 22 (12). Furthermore, these data may also reflect the property of NF-κB to control transcription primarily through distal enhancers or simply indicate that promoter elements are not well defined on a genome-wide basis or that the majority of the NF-κB-bound sites in vivo do not correlate with gene expression of the nearby genes (see below).

**In Vivo NF-κB Binding Sequence Preference Reveals a Novel κB Element**—We examined our data set to determine whether the NF-κB consensus motif, 5′-GGGRRNYCC-3′, is enriched in the isolated p65-bound genomic sequences. Fig. 1B shows that only 6.5% of the identified κB sites fit to the NF-κB consensus, whereas 38% of the bound DNA fragments lack a recognizable NF-κB site (i.e. they bear more than three mismatches from the consensus κB site). We scanned the DNA

---

4 A. Antonaki, C. Demetriades, A. Polyzos, A. Banos, G. Vatsellas, M. D. Lavigne, E. Apostolou, E. Mantouvalou, D. Papadopoulou, G. Mosialos, and D. Thanos, unpublished data.
sequence of the 366 clones for κB-binding sites, and this search revealed an enrichment of the sequence 5′-GGGR(T/A)TYY(C/A)Y-3′ (Fig. 1D). This sequence is similar to the current NF-κB consensus motif (5′-GGGRNNYYCC-3′) differing at only one nucleotide (underlined), an observation that underscores the robustness of our ChIP-cloning approach to identify in an unbiased manner genome-wide binding sites for NF-κB proteins. Of note, we discovered the frequent presence of a T instead of N at position 6 of the consensus (Fig. 1D).

Remarkably, we discovered that 11% of our κB sites contained the specific GGTTTCCACC κB sequence (Fig. 1B) that deviates from the consensus at two nucleotides (underlined), thus increasing the total percentage of sequences bearing two mismatches to 35% (24 + 11%). The GGTTTCCACC DNA site (named the Alu-κB site from this point onwards) is present in specialized Alu-repetitive elements (38, 39). We did not identify Alu-κB sites outside of Alu repeats. Thus, the Alu-κB site is not just another NF-κB DNA-binding site but the most frequent human DNA site on which NF-κB binds in vivo. None of the previous studies identified the Alu-κB site, probably because of the inherent bias of the methods used (10, 12, 14, 36). We have previously identified three Alu-κB sites implicated in

FIGURE 1. Statistical analysis of p65-bound genomic loci. A, the 366 clones were classified by locations relative to the nearest neighbor gene. The pie chart shows the distribution of p65-binding sites located in promoters (−10 kb to the transcription start site), exons, introns, 5′- and 3′-untranslated regions, remote sites (−100 to −100 kb), and intergenic sites (>100 kb away from genes). Putative NF-κB-binding sites were detected using the TESS and TRANSFAC 7.0 platform. B, pie chart illustrating the number of mismatches of the identified κB sites relative to the consensus κB site. The GGTTTCCACC element, although it bears two mismatches, was classified separately. C, shown is the frequency of the GGTTTCCACC element in human and mouse chromosomes. The distribution of the mutant control sequence GGTTATCACC is also shown for both mouse and human chromosomes. D, position weight matrix of the κB sites derived from the ChIP cloning experiments using THEME.

A Novel NF-κB-binding Site in Alu-repetitive Elements

NOVEMBER 4, 2011• VOLUME 286 • NUMBER 44
JOURNAL OF BIOLOGICAL CHEMISTRY 38771
A Novel NF-κB-binding Site in Alu-repetitive Elements

![Diagram](image)

**FIGURE 2. Genome-wide interaction of NF-κB to selected Alu repeats.** A, alignment of 30-bp-long insertions of Alu repeat element family sequences flanking the Alu-κB site. Sequences were retrieved from the NCBI data base and aligned by the Clustal algorithm. The Alu-κB element (5'-GGGTTCACC-3') is boxed. B, chromatin prepared from HeLa cells that were either mock-infected (0 h) or infected for 6 h with Sendai virus was immunoprecipitated with the anti-p65 antibody. The presence of each Alu repeat family in the immunoprecipitated material was detected by PCR. C, pie chart illustrating the distribution of the Alu-κB elements found in our clones located in promoters (>10 kb to the transcription start site), exons, introns, 5' and 3' untranslated regions, remote sites (>100 kb away from genes).

The NF-κB-dependent virus-induced expression of the human IFN-β gene via interchromosomal interactions with the IFN-β locus (22).

Bioinformatics analysis indicated that the GGGTTTCACC κB site occurs 170,755 times in the human genome, whereas only 1,425 copies exist in the mouse genome, a result consistent with the absence of Alu-repetitive elements in mice (Fig. 1C). The Alu-κB site is present in the AluSx, AluSg, and AluY elements and not in the other types of Alu repeats, such as AluSq, AluJo, and AluJb (Fig. 2A). To verify in vivo binding of NF-κB to some of the κB site-containing specific Alu repeats (AluSx, AluSg, and AluY), we carried out ChIP experiments using anti-p65-immunoprecipitated chromatin prepared from mock- and Sendai virus-infected HeLa cells and PCR primers flanking each one of the Alu-κB elements (Fig. 2B (lanes 1 and 2) and 2C). The Alu-κB site is present in the AluSx, AluSg, and AluY and not with AluSq, AluJo, and AluJb repeats lacking the Alu-κB site. Fig. 1C shows the distribution of the GGGTTTCACC Alu-κB element in the human chromosomes in parallel with the distribution of the GGGTTCACC sequence bearing a single base substitution (underlined), used as a control in the bioinformatics analysis. We found that the frequency of the GGGTTTCACC element follows the distribution of the Alu repeats into the human genome, and it depends on the size of chromosomes, thus implying no preference for specific chromosomes (Fig. 1C). Furthermore, this comparison verified the presence of the GGGTTTCACC motif in repetitive elements and that the mutant control sequence occurs in the human genome at the expected abundance for any random 10-mer (2,634 times versus the theoretically expected 3,000 times for the 3 × 10^9-bp human genome). Finally, the distribution of the Alu-κB-containing loci in the genome relative to known genes revealed no preference for specific genomic positions, thus correlating with the overall distribution of Alu elements (Fig. 2C). The Alu-κB site was also highly represented among the DNA fragments that were identified by chromatin immunoprecipitation and cloning of p52/NF-κB2-binding sites in the genome of latently infected human B lymphocytes with EBV (data not shown).

To examine whether our conclusions reflect the global general binding properties of NF-κB and are not the result of the relatively small data sets of the ChIP-cloning experiment, we took advantage of the public availability of recent raw data of a ChIP-seq experiment examining p65 binding to the human genome following TNF-α activation in the lymphoblastoid cell line GM12878 from the ENCODE consortium (available from the University of California Santa Cruz Web site). The data were generated by the Snyder, Gerstein, and Weissman laboratories (Yale University), the Farnham laboratory (University of California, Davis), and the Struhl laboratory (Harvard University). Our analysis of these p65-ChIP-seq data is shown in supplementary Fig. 2, and the procedures for the analysis are described under “Experimental Procedures.” In a remarkable agreement with our ChIP-cloning analysis (Fig. 1), we found that only a small fraction (17.8%) of the total NF-κB binding events occur at promoter regions (<10 kb from the start site of transcription), a result in agreement with a recent publication examining the relationship of NF-κB and Bcl-6 binding in the mouse genome in which 10% only of the p65 bound regions were found at promoters (37). The distribution of NF-κB between intergenic regions, introns, exons, and 5’- and 3’-UTRs is very similar to our ChIP-cloning experiments (compare Fig. 1A with supplementary Fig. 2A). Furthermore, the consensus NF-κB DNA binding motif derived from the ChIP-cloning analysis is similar to the consensus motif derived form the
A Novel NF-κB-binding Site in Alu-repetitive Elements

analysis of the ENCODE-derived ChIP-seq data (compare Fig. 1D with supplemental Fig. 2B). The small differences in the consensus sequences are most likely due to the inherent bias of the ChIP-seq analysis toward the higher affinity binding sites (the ChIP-seq-derived motif is more similar to the consensus κB site than the consensus derived from the ChIP-cloning data). This is consistent with previous biochemical studies using permutations of the κB motif in in vitro binding assays (11). Finally, we discovered that a significant number of peaks do not contain a recognizable NF-κB binding motif (data not shown). More specifically, 67% of all ChIP-seq peaks contain sequences that deviate from the consensus by more than two mismatches, a result in agreement with our ChIP-cloning analysis (73%).

Importantly, we found a significant percentage of p65 peaks that coincides with Alu repeats at frequencies very similar to the ones we described with the ChIP-cloning approach (for a comparison, see supplemental Fig. 2). Taken together, these data strongly suggest that our major finding of an extensive interaction of NF-κB with Alu repeats in the human genome is not the result of the smaller data sets used and analyzed by the ChIP-cloning method. In summary, our combined ChIP-cloning and ChIP-seq analyses revealed the existence of the previously unknown Alu-κB site, which is the most frequent binding site targeted by NF-κB in human cells, and all Alu-κB sites are part of Alu repeats.

Recruitment of NF-κB Proteins to Multiple Genomic DNA Sites in Vivo and in Vitro—To examine DNA binding of NF-κB and local histone modifications to the wide variety of identified putative target sites, we carried out chromatin immunoprecipitation experiments using pairs of primers corresponding to a randomly selected set of clones bearing different κB motifs. Fig. 3A shows that all selected genomic clones are bound (with varying affinities) in vivo by the p65 subunit of NF-κB in a virus infection-inducible manner (lanes 1 and 2), a result supporting the validity and specificity of our ChIP-cloning approach. Thus, NF-κB binds to a wide variety of DNA sites, located at many different positions relative to genes, including DNA binding to the GGGTTTACC Alu-κB element. Interestingly, a similar pattern of NF-κB association was obtained when the cells were treated with TNF-α to induce the NF-κB DNA binding activity (Fig. 3A, lanes 7 and 8). Importantly, we tested three of our clones (clones 48, 138, and 224) lacking a recognizable κB site (supplemental Fig. 1) and verified that they are indeed bound by NF-κB, thus supporting further our initial conclusion that NF-κB is recruited to genomic sites lacking κB elements. A subset of the NF-κB-bound loci can also associate with HMGI(Y) (Fig. 3A), suggesting a function of the protein in facilitating NF-κB DNA binding activity at genomic loci scattered throughout the genome.

To examine whether different NF-κB proteins could occupy the same genomic sites, we carried out side-by-side chromatin immunoprecipitation experiments using antibodies against the p65, p52, and RelB proteins and chromatin prepared from latently infected human B cells with EBV, in which different heterodimeric forms of NF-κB proteins are constitutively nuclear. As seen in Fig. 3B, we found that all tested binding events occurred at all sites that can be occupied by all three NF-κB family members. Thus, in vivo, κB sites are occupied by multiple NF-κB family members stimulated by different inducers. This REL protein exchange is also consistent with a report examining the Rel subunit composition on the promoters of five NF-κB-regulated genes during dendritic cell maturation (13).

To test whether binding of NF-κB at genomic sites located away from known enhancers or promoters can induce local chromatin modifications (e.g. acetylation) and the recruitment of general transcription factors, we carried out ChIP experiments using antibodies against acH4, acH3, TBP, and PolII. Fig. 4 shows that in ~50% of the tested genomic sites, the local chromatin is either preacetylated at various levels (clones 8, 31, 77, 87, 95, 175, 366, 315, 216, 223, 217, 307, and 310), or it becomes further acetylated upon NF-κB binding (clones 11, 77, 87, 313, and 366), or the acetylation marks are decreased upon virus infection and NF-κB binding (probably due to nucleosome loss). These histone acetylation patterns do not correlate with the presence or absence of Alu repeats (Fig. 4). Furthermore, increased histone acetylation in these loci does not always correlate with transcriptional activation in response to virus infection (see Figs. 6 and 7A) and/or the recruitment of TBP and PolII (Fig. 4). In summary, NF-κB can induce distinct histone acetylation patterns and recruitment programs upon binding to different Alu-κB sites, thus underscoring the role of the κB context in NF-κB function. It should be mentioned, however, that the observed chromatin modifications and recruitment of general transcription factors could be actually occurring at remote sites with which these κB elements physically associate via DNA looping or interchromosomal interactions and are captured by the ChIP assay (22, 40, 41). Finally, these results are in agreement with a study in yeast showing that binding of GCN4 to non-promoter regions can induce local histone acetylation and recruitment of basal transcription factors (42).

The ability of NF-κB to interact with the Alu-κB site was evaluated by carrying out electrophoretic mobility shift assays (EMSAs) using oligonucleotides encompassing Alu-κB sites taken from various clones and nuclear extracts from mock-, virus-, and TNF-α-treated cells or recombinant NF-κB proteins (p50/p65 heterodimer). Fig. 5A shows that nuclear activated (lanes 1–20) or recombinant NF-κB (lanes 21–30) binds to the Alu-κB element derived from different genomic sites. Interestingly, the relative affinities of the in vitro DNA binding experiments are in agreement with the in vivo occupancy data presented in Fig. 3. For example, clone 77 is a weak binding site for NF-κB both in vivo and in vitro (compare Fig. 3A with Fig. 5A). These DNA-protein interactions are specific because they can be competed away by the WT and not by the mutant oligonucleotides (Fig. 5C, compare lane 2 with lanes 3–6). The protein–DNA complexes formed in nuclear extracts prepared from TNF-α-treated cells on the Alu-κB element contain the p50 and p65 subunits of NF-κB because the complexes were abrogated by the addition of the corresponding antibodies (Fig. 5C).

Finally, the titration competition experiment of Fig. 5D suggests that the affinity of NF-κB for the Alu-κB site is ~5 times less than the affinity for the typical PRDII5 site taken from the IFN-β gene.

5 The abbreviation used is: PRDII, positive regulatory domain II.
Identification of Virus Infection-inducible Genes Regulated by NF-κB—To identify p65-bound genes that are responsive to virus infection, we carried out transcriptomics analysis in HeLa cells using the Affymetrix platform to identify genes that are differentially expressed after virus infection for 6 and 12 h. We identified 1,262 genes that are significantly up-regulated upon virus infection (supplemental Fig. 3). Of these 1,262 virus-inducible genes, we found that only 36 are located near the p65-bound loci identified by the ChIP-cloning approach (Fig. 6A and supplemental Fig. 4), and only 64 virus-inducible genes represent previously known and well characterized NF-κB targets (Fig. 6B and supplemental Fig. 5). Interestingly, of the 462
A Novel NF-κB-binding Site in Alu-repetitive Elements

previously known NF-κB targets, only 11 were identified in our ChIP-cloning assay (Fig. 6C and supplemental Fig. 4). Thus, the 36 novel virus-inducible genes that are associated with our list of identified sequences represent a reliable list of NF-κB target genes in virus-infected human cells. The 36 genes we identified as novel putative direct NF-κB targets encode transcription factors (ZFR, ZNF704, and PSGF5), signaling proteins (CDKN2C, NF-κB), and/or validated NF-κB targets (ZFR, ZNF704, and PSGF5), signaling proteins (CDKN2C, NF-κB, and/or validated NF-κB targets), and only 64 of the previously known 462 verified virus targets (Fig. 6D, right). Thus, a random sampling of p65-bound genomic loci correlates with gene expression only at a 10% frequency (36 of 366), and none of these gene expression patterns correlates with the magnitude of the transcriptional activation in response to virus infection. Taken together these data strongly suggest that most of the NF-κB-bound genomic sites are not capable of promoting transcriptional activation, and/or NF-κB is in general insufficient to trigger gene expression.

Transcriptional Activation Properties of Genomic NF-κB Sites—To investigate the transcriptional properties of the novel NF-κB sites, we generated reporter constructs in which randomly selected genomic fragments bearing the κB elements were placed upstream or downstream from the thymidine kinase promoter linked to the luciferase gene. These reporters were cotransfected into HeLa cells along with an empty vector or a vector expressing the p65 subunit of NF-κB, and the luciferase activity was determined in mock- and virus-infected cells. Remarkably, we found that only 30% (7 of 21) of the tested elements activated the luciferase reporter in response to p65 overexpression, and of these, only two (clones 77 and 216) could mediate Sendai virus-induced transactivation of the luciferase reporter (Fig. 7A). Western blotting experiments verified that the p65 protein was expressed at similar levels in all samples (Fig. 7B). Interestingly, the majority of the genomic sites bearing the Alu-κB elements did not respond to virus infection or p65-dependent transactivation except the clones 15, 16, 23, and 71, whereas the non Alu-κB clone 216 was activated at the highest levels by overexpressed p65 (Fig. 7A). We do not yet know why clone 216 functions as a strong p65 responder, but we can assume that high level p65 expression could lead to binding at very low affinity non-consensus sites or to the indirect recruitment of p65 to the DNA by unknown transcription factors. Remarkably, we found that only 30% (7 of 21) of the tested elements activated the luciferase reporter in response to p65 overexpression, and of these, only two (clones 77 and 216) could mediate Sendai virus-induced transactivation of the luciferase reporter (Fig. 7A). Western blotting experiments verified that the p65 protein was expressed at similar levels in all samples (Fig. 7B). Interestingly, the majority of the genomic sites bearing the Alu-κB elements did not respond to virus infection or p65-dependent transactivation except the clones 15, 16, 23, and 71, whereas the non Alu-κB clone 216 was activated at the highest levels by overexpressed p65 (Fig. 7A). We do not yet know why clone 216 functions as a strong p65 responder, but we can assume that high level p65 expression could lead to binding at very low affinity non-consensus sites or to the indirect recruitment of p65 to the DNA by unknown transcription factors bound to clone 216. Surprisingly, the experiment of Fig. 7C shows that oligomerization of the Alu-κB element only responded weakly to high level expression of p65. More specifically, the amounts of activated transcription were ~30 times less compared with the ones obtained with the PRDII NF-κB site taken from the IFN-β enhancer/promoter, under conditions of saturating NF-κB DNA binding (also see below).

To determine why NF-κB bound to the Alu-κB element fails to promote high levels of transcription, we carried out in vitro transcription experiments allowing either single or multiple rounds of transcription reinitiation using recombinant NF-κB (p50/p65 heterodimer) and HeLa nuclear extracts. Our DNA templates contained four copies of the PRDII element (GGGAAATTC) or four copies of the Alu-κB element (GGGTTCACCA) taken from two different genomic clones (bearing different flanking sequences) placed immediately upstream of the IFN-β core promoter (33). The DNA templates were incubated with saturating amounts of recombinant NF-κB, followed by incubation with HeLa nuclear extracts for 30 min to allow the assembly of preinitiation complexes. Next, NTPs were added, followed by the addition of Sarkosyl, which
prevents the reassembly of initiation complexes, or buffer, and the samples were incubated for 30 min (33). Correctly initiated transcripts were detected by primer extension. Fig. 7D shows that when multiple rounds of transcription were allowed, NF-kB activated transcription 36-fold from the PRDII site and only 7-fold from each of the Alu-kB sites (compare lane 2 with lanes 6 and 10). These results are in general agreement with the in vivo transfection experiments of Fig. 7C. Remarkably, in the presence of Sarkosyl, NF-kB activated transcription ~4-fold from both the PRDII and the Alu-kB sites, implying that the PRDII element supports ~9 rounds of transcription from PRDII and ~1.7 rounds from the Alu-kB element. Thus, the distinct transcriptional strength of NF-kB on PRDII versus the Alu-kB site is due to the inability of NF-kB to promote multiple transcriptions.

FIGURE 5. Interaction of NF-kB with the Alu-kB element in vitro. A, an EMSA using radiolabeled oligonucleotide probes derived from randomly selected clones and nuclear extracts from untreated (TNF) or TNFα-treated (TNFα) HeLa cells (lanes 1–10) or recombinant NF-kB (p50/p65) (lanes 11–20). B, sequence of the probes and the competitor oligonucleotides used in the EMSA experiments. C, EMSA supershift experiments, using the Alu-kB clone 31 as a probe and nuclear extracts prepared from TNF-α-treated HeLa cells. The specificity of the NF-kB-containing complexes was verified by competition (lanes 3–6) and by incubation of the complexes with the anti-p65 and anti-p50 antibody (lanes 7–8). The NF-kB-containing complexes are labeled as a, and the supershifted complexes are labeled as c and b. D, shown is an EMSA experiment using the Alu-kB element derived from clone 11 as a probe and recombinant NF-kB (p50/p65 heterodimer). The DNA-binding reactions were carried out in the absence (lane 1) or in the presence of cold PRDII (lanes 3–8) or Alu-kB clone 11 competitors at the indicated molar ratios.
Distinct Functional Properties of Alu-/H9260B Elements—We have previously cloned three Alu-/H9260B genomic sites by virtue of their ability to physically associate through NF-/H9260B with the IFN-/H9252 gene via interchromosomal interactions (22). These associations are required for the delivery of Alu-/H9260B-bound NF-/H9260B to the IFN-/H9252 gene and are essential for virus-induced IFN-/H9252 expression.

FIGURE 6. Comparison of genome-wide p65 chromatin occupancy with virus-induced gene expression. A, comparison of the 1,262 genes induced by virus infection with the 374 genes located near the 366 p65-bound loci as displayed in Venn diagrams. The number of genes identified in each case is indicated. B, Venn diagram comparing the number of virus-inducible genes identified in our study with all previously known NF-/H9260B target genes obtained from the Boston University NF-/H9260B Transcription Factors Web site. C, Venn diagram comparing all known NF-/H9260B target genes (462 genes) with the 374 genes located near the 366 p65-bound loci identified by our ChIP-cloning approach. D, heat map showing the expression profile of virus-induced genes at 6 and 12 h postinfection. Red or black represents increase or no change in gene expression relative to uninfected cells, respectively. Only genes displaying significant changes in expression levels (>1.5-fold or 1/1.5-fold; false discovery rate <5%) as compared with the uninfected control in three biological replicates are shown. The moving average of genes in a window of 60 genes that are known to be regulated by NF-/H9260B is shown on the left. The right part illustrates a similar analysis performed with the 64 and 36 genes identified by our study to contain an NF-/H9260B site and to be activated by virus infection. The straight line represents the expected percentage of NF-/H9260B-regulated genes. The group of 64 genes corresponds to previously known NF-/H9260B targets that were identified in our DNA microarray analysis to be induced by virus infection, whereas the group of 36 genes corresponds to genes identified by our ChIP cloning analysis and shown to be induced by virus infection.

rounds of reinitiation of transcription from the Alu-/H9260B element.
expression. Because the Alu-κB site is incapable of supporting by itself high levels of transcription, we hypothesized that the Alu-κB elements may play a more general role in interchromosomal interactions. We have tested three randomly picked clones derived from the ChIP-cloning library (clones 11, 17, and 39) for their ability to interact with the IFN-β gene side-by-side with the previously characterized 4C-ChIP clones (22). Here we show representative results for clone 11, but similar results were obtained with clones 17 and 39. Fig. 8A shows the maximum alignment of the two Alu repeat sequences surrounding

FIGURE 7. Transcriptional properties of genomic p65-bound loci. A, HeLa cells were cotransfected with reporter constructs bearing randomly selected p65-bound loci upstream of the thymidine kinase (TK) minimal promoter directing the expression of luciferase with or without a p65 expression vector. The cells were either mock- or virus-infected for 12 h, and the luciferase activity was determined and plotted after normalization with lacZ. B, Western blotting using some of the extracts generated in Fig. 7A reacted with the anti-p65 antibody (sc-109, Santa Cruz Biotechnology, Inc.). An anti-β-actin antibody (sc-1616, Santa Cruz Biotechnology, Inc.) was used to detect β-actin as an internal control. C, HeLa cells were cotransfected with reporter constructs bearing one, two, or four copies of the GGGTTACC or PRDII (GGGAATTCC) element cloned immediately upstream of the IFN-β minimal promoter in the presence or the absence of the p65 expression vector. The cells were either mock- or virus-infected for 12 h before the luciferase activity was determined. D, shown is an in vitro transcription assay using HeLa nuclear extracts and templates bearing four copies of the PRDII element (PRDIIx4LUC) or four copies of the Alu-κB element (Alu-κB4xLUC). The sequence of the κB oligonucleotides used to generate the templates is shown in Fig. 5B. The in vitro transcription reactions were carried out in the absence or the presence of saturated amounts of recombinant NF-κB according to the scheme shown at the left of the gel. The NF-κB-dependent transcripts generated in the absence or presence of Sarkosyl (an inhibitor of the reinitiation of transcription) are indicated by an arrow (top) and were quantified by the ImageQuant software (Molecular Dynamics) to calculate the rounds of transcription that took place (bottom). A representative experiment of two is shown. Error bars, S.D.
the Alu-κB site from the ChIP-cloning clone 11 and the 4C-ChIP clone 21 shown previously to interact with the IFN-β locus in an NF-κB-dependent manner. As seen in Fig. 8, the two Alu-κB clones share 80% identity. To determine whether binding of NF-κB to different Alu-κB clones can increase IFN-β gene expression by promoting interchromosomal interactions, we carried out three-dimensional DNA FISH analysis using HeLa cells that were either mock- or virus-infected for 4 h. The IFN-β bacterial artificial chromosome probe (green) detects 6 or 7 IFN-β genes, and the clone 21 bacterial artificial chromosome probe detects three copies of clone 21, because HeLa cells are polyploid. In agreement with our previous work (22), virus infection induces interchromosomal interactions between a single allele of clone 21 and one allele of IFN-β in ~9% of the

FIGURE 8. Distinct functional properties of Alu-κB repeat elements. A, maximum DNA sequence alignment between the chip-cloning clone 11 and the 4C-ChIP clone 21. Identical sequences are boxed, and the Alu-κB element is indicated. B, DNA FISH analysis examining interchromosomal interactions between IFN-β and Alu-κB clones. A digoxigenin-labeled bacterial artificial chromosome DNA probe was used to detect the endogenous IFN-β locus and was visualized with FITC-conjugated anti-DIG antibody (green). Biotin-labeled bacterial artificial chromosome DNA probes were used to detect the endogenous Alu-κB clone 21 and 11 loci, as indicated in the figure (top), and visualized with Alexa 568-conjugated streptavidin (red). The stably integrated plasmids bearing clones 11 and 21 were detected through the preparation of plasmid DNA (pKS) biotin-labeled probes (bottom). The circle indicates the colocalized alleles observed only in virus-infected cells. The bottom of B shows a diagrammatic representation of the percentage (mean ± S.D. (error bars)) of cells with colocalization between IFN-β and each of the Alu-κB clones in mock- and virus-infected cells. C, HeLa cells stably transfected with Bluescript-based constructs bearing the 4C-ChIP clone 21 and the ChIP-cloning clone 11 along with untransfected HeLa cells or HeLa cells bearing the empty vector were either mock-infected or infected with Sendai virus for 4, 6, or 8 h, and the isolated RNAs were used as templates for RT-PCR analysis using primers specific for IFN-β. The PCR products were detected by gel electrophoresis.
cells. In sharp contrast, no such interaction was detected between the ChIP-cloning clone 11 and the IFN-\(\beta\) locus (Fig. 8B). To investigate whether clone 11 could potentiate IFN-\(\beta\) expression in the absence of physical interactions with the IFN-\(\beta\) locus, we generated stable HeLa cells bearing multiple copies (35–40) of clones 21 and 11. Both stable cell lines bear the inserted DNA at a single integration site (see below). The stable cell lines bearing the empty vector and the Alu-\(\kappa\)B clones 11 and 21 were infected with virus for different amounts of time, and IFN-\(\beta\) gene expression was determined by RT-PCR. Fig. 8C shows that the cells bearing multiple copies of clone 21 express the IFN-\(\beta\) transcript at much higher levels compared with control cells (compare lanes 1–8 with 13–16). Interestingly, no effect on IFN-\(\beta\) expression was detected in the cells bearing multiple copies of clone 11. The increased levels of IFN-\(\beta\) expression in clone 21 cells correlate with interchromosomal interactions between the IFN-\(\beta\) gene and clone 21, although clone 21 has been inserted at a different site from its natural position. Interestingly, no such interactions were observed in cells bearing clone 11. We conclude that different Alu-\(\kappa\)B repeats are functionally distinct despite their significant DNA homology. It is important to mention that different Alu-\(\kappa\)B repeats may interact with different target genes. Furthermore, these data also demonstrate that the property of some Alu-\(\kappa\)B repeats to participate in interchromosomal interactions is not determined by their genomic context as revealed by the ability of the stable clone 21 to interact with the IFN-\(\beta\) gene.

**DISCUSSION**

Here we have addressed the interaction of the transcription factor NF-\(\kappa\)B with the human genome and found that \(\sim 11\%\) of the total NF-\(\kappa\)B DNA binding events occur at a specialized type of Alu repeat \(\kappa\)B element that could function by controlling the activity of NF-\(\kappa\)B. Furthermore, we demonstrate that the vast majority of the NF-\(\kappa\)B DNA binding events do not directly correlate with gene regulation, a conclusion in agreement with earlier preliminary observations (10, 12, 14). All tested novel NF-\(\kappa\)B-bound genomic elements are true \(\kappa\)B-binding sites despite the fact that many of them are lacking a recognizable \(\kappa\)B motif or a \(\kappa\)B consensus. NF-\(\kappa\)B may interact with these elements via another DNA-binding protein, and/or these loci may interact with distant NF-\(\kappa\)B sites via chromatin looping or interchromosomal interactions.

The ChIP-cloning approach provided novel information regarding the genome-wide distribution of NF-\(\kappa\)B (10, 12, 14, 37). The main advantage of our ChIP-cloning approach is the identification in an unbiased manner of NF-\(\kappa\)B-binding DNA sites of variable affinities (low, medium, and high), whereas in all previous studies, most of the identified \(\kappa\)B sites were of relatively high affinity. The second advantage is that in all high throughput methods (ChIP-cloning PET, ChIP-chip, and ChIP-seq) the analysis relies on the use of specific algorithms either to identify the peaks of tagged sequences and/or map the short sequences derived from high throughput sequencing to the genome or to measure the relative hybridization intensity. It is not surprising that different algorithms can produce different results using the same set of data. In sharp contrast, although our approach is of medium scale, we have manually curated the sequencing data, thus increasing the confidence of the results. Finally, high throughput approaches based on hybridization (ChIP-chip) and sequencing (ChIP-seq) may mask putative DNA transcription factor-binding sites at repetitive sequences. This is the main reason that the Alu-\(\kappa\)B site remained elusive until we carried out the ChIP-cloning experiment. On the other hand, the disadvantage of our ChIP-cloning approach is the identification of only a small number of binding sites when compared with the other high throughput approaches.

The fact that of the 100 NF-\(\kappa\)B-bound genomic loci identified in our analysis, only a few correlate with changes in expression of the associated genes suggests that most NF-\(\kappa\)B binding events may not have functional consequences under the specific conditions tested here (virus infection), but they could be functional under another condition. These conclusions are also in agreement with a recent study examining genome-wide binding of liver transcription factors in five different vertebrate species (43). Thus, these unexpected findings raise the question of why there would be so many “nonfunctional” binding sites. We can imagine the following possibilities. 1) The assignment of the binding site to the nearby gene may not be accurate because the bound factor may regulate the expression of a different gene via the formation of long range cis loops or interchromosomal interactions (reviewed in Ref. 46). 2) Binding sites located away from genes may correspond to docking DNA elements used by remote enhancers to reach their target promoters through the progressive formation of transient loops along the DNA (41). 3) Transcription factors may bind to these “irrelevant” sites transiently only (44), suggesting that the “snapshot” of transcription factor DNA binding taken by ChIP-based approaches most likely reflects factors on the “move” (i.e. the continuous association and dissociation from non-relevant sites until they encounter a promoter). 4) Some NF-\(\kappa\)B-binding sites may be used as a shuttle receiving NF-\(\kappa\)B and “delivering” it to target genes via DNA looping and/or interchromosomal interactions. This notion is supported by our recent discovery of three distally located Alu-\(\kappa\)B elements promoting NF-\(\kappa\)B-dependent expression of the interferon-\(\beta\) gene during the early phase of virus infection (22). Specifically, upon nuclear translocation, NF-\(\kappa\)B initially associates with three Alu-\(\kappa\)B elements and subsequently is transferred to the IFN-\(\beta\) gene via interchromosomal interactions, resulting in the initiation of transcription.

The main question arising from our study regards the role of NF-\(\kappa\)B when bound to the Alu-\(\kappa\)B sites. The human genome contains over one million Alu repeat elements (reviewed in Refs. 23 and 24), and \(\sim 20\%\) of those contain the non-consensus Alu-\(\kappa\)B site GGGTTCACC. Our ChIP-cloning and ChIP-seq data showed that the Alu-\(\kappa\)B site is the most frequently recognized \(\kappa\)B element in the human genome because \(\sim 11\%\) of the total NF-\(\kappa\)B DNA binding events occur at the Alu-\(\kappa\)B site. The Alu-\(\kappa\)B site is distinct from all other \(\kappa\)B sites because it cannot support direct transcriptional activation due to its inability to promote reinitiation of transcription. We speculate that this is due to allosteric changes induced on the Rel proteins by the specific Alu-\(\kappa\)B DNA sequence, a hypothesis consistent with previous studies demonstrating DNA site-induced conforma-
tional changes on DNA binding proteins, thus affecting their ability to activate or repress transcription (27, 45). However, depending on the context and the nearby transcription factors, the Alu-κB element can work as a virus-inducible enhancer (Fig. 7A). Is there a general function for the Alu-κB sites, or could they have different functions depending on their genomic context, cell type, environmental cues, developmental stage, etc.? We favor the latter possibility because we showed that, for example, their ability to participate in interchromosomal interactions is not general and that some of them, in the appropriate context, can work as enhancers. Below we summarize some possible functions, and we emphasize that our current technical inability to define the whole spectrum of specific roles for the Alu-κB elements does not necessarily mean that they are without function. We hypothesize the following. First, based on the fact that proteins in solution exist as a population of different conformers (46), we propose that nuclear NF-κB could exist in many conformers, some of which are transcriptionally inert but can bind to the Alu-κB site. Thus, binding of this specific NF-κB conformer to the Alu-κB site could lead to its dramatic reduction in the population of the NF-κB ensembles, thus allowing the rest of the conformers, which are transcriptionally competent, to bind and activate target genes. According to this model, the Alu-κB site functions as a sink “removing” transcriptionally inert NF-κB molecules, with the subsequent optimization of the interactions taking place via induced fit and/or DNA-induced allostery.

Second, Alu repeats have been found to contain several potential functional transcription factor-binding sites that could modulate gene expression through their presence in enhancers, promoters, and/or introns (34). In agreement with this idea is the recent observation that repetitive DNA elements spread throughout the genome, including regulatory regions, contain bona fide DNA sites for transcription factors (OCT4, NANOG, and CTCF), thus engaging the associated genes in the core regulatory network of stem cells (25). This study, together with the results presented herein, suggests that binding of transcription factors to repetitive DNA elements could be associated with regulatory expansions, thus playing an important role in expanding the repertoire of binding sites (see also Ref. 47). Third, Alu-κB elements could also act as transcriptionally inert docking sites by sequestering the NF-κB subunits immediately after nuclear entry, thus preventing excessive targeting and superactivation of promoters. This is supported by our finding that the Alu-κB element is nucleosome-free and therefore fully accessible for factor interactions (not shown).

The wide distribution of the Alu-κB elements combined with other nearby cis-acting elements may expand the NF-κB activity by engaging more genes into specific regulatory networks. This level of NF-κB regulation may be a characteristic feature of primate phylogeny because the Alu-κB element is not present in other species. Furthermore, the larger number (~3 times) of species-specific insertions of Alu-repetitive elements in the human genome, compared with the chimpanzee genome, suggests a mechanism for differential regulation of NF-κB, even between primates (7), thus possibly contributing to human evolution. In addition, the high density of Alu elements in the genome, when taken together with their direct involvement in interchromosomal interactions (2), suggests that Alu repeats together with their bound transcription factors could play an important role in recombination events triggering genomic instability and/or genetic disorders (8). The latter point is further supported by a recent study showing that genome-wide binding of NF-κB differs by 25% between individuals, and most of these differences are localized in intergenic regions and are due to SNPs (48). Remarkably, and in agreement with data presented herein, most of these binding differences between individuals are not associated with differences in gene expression. In summary, these findings, combined with our data and the observation that Alu repeats are polymorphic in nature, support the idea that differences in gene expression between individuals and between humans and other primates could be due to long range effects of the polymorphic Alu repeats and/or other genomic sites like some of those identified in this study.

Acknowledgments—We thank Arg Efstratiadis, Ethan Ford, Menie Merika, George Panayotou, and George Theirotos for critical reading of the manuscript.

REFERENCES

1. Vallabhapurapu, S., and Karin, M. (2009) Annu. Rev. Immunol. 27, 693–733
2. Hayden, M. S., and Ghosh, S. (2008) Cell 132, 344–362
3. Li, Q., and Verma, I. M. (2002) Nat. Rev. Immunol. 2, 725–734
4. Fan, W., and Lenardo, M. J. (2009) Cold Spring Harb. Perspect. Biol. 1, a00067
5. Bonizzi, G., and Karin, M. (2004) Trends Immunol. 25, 280–288
6. Pasparakis, M. (2009) Nat. Rev. Immunol. 9, 778–788
7. Hoffmann, A., Natoli, G., and Ghosh, G. (2006) Oncogene 25, 6706–6716
8. Natoli, G., Saccani, S., Bossisio, D., and Marazzi, I. (2005) Nat. Immunol. 6, 439–445
9. Chen, F. E., and Ghosh, G. (1999) Oncogene 18, 6845–6852
10. Lim, C. A., Yao, F., Wong, J. I., George, J., Xu, H., Chiu, K. P., Sung, W. K., Lipovich, L., Vega, V. B., Chen, J., Shabah, A., Zhao, X. D., Hibberd, M., Wei, C. L., Lim, B., Ng, H. H., Ruan, Y., and Chin, K. C. (2007) Mol. Cell 27, 622–635
11. Linnell, J., Mott, R., Field, S., Kwiatkowski, D. P., Ragoussis, I., Udalova, I. A. (2004) Nucleic Acids Res. 32, e44
12. Martone, R., Euskirchen, G., Bertone, P., Hartman, S., Royce, T. E., Luscombe, N. M., Rinn, J. L., Nelson, F. K., Miller, P., Gerstein, M., Weissman, S., and Snyder, M. (2003) Proc. Natl. Acad. Sci. U.S.A. 100, 12247–12252
13. Saccani, S., Pantano, S., and Natoli, G. (2003) Mol. Cell 11, 1563–1574
14. Schreiber, J., Jenner, R. G., Murray, H. L., Gerber, G. K., Gifford, D. K., and Young, R. A. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 5899–5904
15. Udalova, I. A., Mott, R., Field, D., and Kwiatkowski, D. (2002) Proc. Natl. Acad. Sci. U.S.A. 99, 8167–8172
16. Farnham, P. J. (2009) Nat. Rev. Genet. 10, 605–616
17. Fujiwara, T., O’Geen, H., Keles, S., Blahnik, K., Linenmann, A. K., Kang, Y. A., Choi, K., Farnham, P. J., and Bresnick, E. H. (2009) Mol. Cell 36, 667–681
18. Hollenhorst, P. C., Chandler, K. J., Poulsen, R. L., Johnson, W. E., Speck, N. A., and Graves, B. J. (2009) PLoS Genet. 5, e1000778
19. Reed, B. D., Charos, A. E., Szekely, A. M., Weissman, S. M., and Snyder, M. (2008) PLoS Genet. 4, e1000133
20. Wodarell, E. D., Bilenky, M., Cullum, R., Thiessen, N., Dapimar, M., Delaney, A., Varhol, R., Zhao, Y., Zeng, T., Bernier, B., Ingham, M., Hirst, M., Robertson, G., Marra, M. A., Jones, S., and Hoodless, P. A. (2008) Nucleic Acids Res. 36, 4549–4564
21. Yu, M., Riva, L., Xie, H., Schindler, Y., Moran, T. B., Cheng, Y., Yu, D., Hardison, R., Weiss, M. J., Orkin, S. H., Bernstein, B. E., Frankel, E., and Cantor, A. B. (2009) Mol. Cell 36, 682–695
A Novel NF-κB-binding Site in Alu-repetitive Elements

22. Apostolou, E., and Thanos, D. (2008) Cell 134, 85–96
23. Cordaux, R., and Batzer, M. A. (2009) Nat. Rev. Genet. 10, 691–703
24. Deininger, P. L., and Batzer, M. A. (1999) Mol. Genet. Metab. 67, 183–193
25. Kunarso, G., Chia, N. Y., Jeyakani, I., Hwang, C., Lu, X., Chan, Y. S, Ng, H. H., and Bourque, G. (2010) Nat. Genet. 42, 631–634
26. Polak, P., and Domany, E. (2006) BMC Genomics 7, 133–147
27. Leung, T. H., Hoffmann, A., and Baltimore, D. (2004) Cell 118, 453–464
28. Merika, M., Williams, A. J, Chen, G., Collins, T., and Thanos, D. (1998) Mol. Cell 1, 277–287
29. Thanos, D., and Maniatis, T. (1992) Cell 71, 777–789
30. Lomvardas, S., and Thanos, D. (2002) Cell 110, 261–271
31. Agelopoulos, M., and Thanos, D. (2006) EMBO J. 25, 4843–4853
32. White, R. B., and Ziman, M.R. (2006) Biochem. Biophys. Res. Commun. 346, 479–483
33. Koutroubas, G., Merika, M., and Thanos, D. (2008) Mol. Cell. Biol. 28, 926–938
34. Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003) Bioinfor-matics 19, 185–193
35. Tusher, V. G., Tibshirani, R., and Chu, G. (2001) Proc. Natl. Acad. Sci. 98, 5116–5121
36. Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., Lempicki, R. A. (2003) Genome Biol. 4, P3
37. Barish, G. D., Yu, R. T., Karunasisri, M., Ocampo, C. B., Dixon, J., Benner, C., Dent, A. L., Tangirala, R. K., and Evans, R. M. (2010) Genes Dev. 24, 2760–2765
38. Häslar, J., Samuelsson, T., and Strub, K. (2007) Cell Mol. Life Sci. 64, 1793–1800
39. Jurka, J. (2004) Curr. Opin. Genet. Dev. 14, 603–608
40. Hatzis, P., and Talianidis, I. (2002) Mol. Cell 10, 1467–1477
41. Nolis, I. K., McKay, D. J., Mantouvalou, E., Lomvardas, S., Merika, M., and Thanos, D. (2009) Proc. Natl. Acad. Sci. U.S.A. 106, 20222–20227
42. Topalidou, I., and Thireos, G. (2003) EMBO Rep. 4, 872–876
43. Schmidt, D., Wilson, M. D., Ballester, B., Schwalie, P. C., Brown, G. D., Marshall, A., Kutter, C., Watt, S., Martinez-limenez, C. P., Mackay, S., Talianidis, I., Flicek, P., Odom, D. T. (2010) Science 328, 1036–1040
44. Hager, G. L., McNally, J. G., and Misteli, T. (2009) Mol. Cell 35, 741–753
45. Beckett D. (2009) Proc. Natl. Acad. Sci. U.S.A. 106, 22035–22036
46. Pan, Y., Tsai C. J., Ma, B., and Nussinov, R. (2010) Trends Genet. 26, 75–83
47. Bourque, G., Leong, B., Vega, V. B., Chen, X., Lee, Y. L., Srinivasan, K. G., Chew, J. L., Ruan, Y., Wei, C. L., Ng, H. H., and Liu, E. T. (2008) Genome Res. 18, 1752–1762
48. Kasowski, M., Grubert, F., Heffelfinger, C., Hariharan, M., Asabere, A., Waszak, S. M., Hабеггер, L., Rozowsky, J., Shi, M., Urban, A. E., Hong, M. Y., Karczewski, K. I., Huber, W., Weissman, S. M., Gerstein, M. B., Korbel, J. O., and Snyder, M. (2010) Science 328, 232–235