MOLECULAR AND CELLULAR BIOLOGY, Apr. 1997, p. 2279–2290 Vol. 17, No. 4
0270-7306/97/$04.00+0


# Regulation of the Myeloid-Cell-Expressed Human gp91-phox Gene As Studied by Transfer of Yeast Artificial Chromosome Clones into Embryonic Stem Cells: Suppression of a Variegated Cellular Pattern of Expression Requires a Full Complement of Distant *cis* Elements

LESTER L. LIEN,[1] YISHENG LEE,[1]† AND STUART H. ORKIN[1,2]*

*Division of Hematology/Oncology, Children's Hospital, Dana-Farber Cancer Institute, and Department of Pediatrics, Harvard Medical School,[1] and Howard Hughes Medical Institute,[2] Boston, Massachusetts 02115*

Received 19 August 1996/Returned for modification 23 October 1996/Accepted 14 January 1997

**Identifying the full repertoire of *cis* elements required for gene expression in mammalian cells (or animals) is challenging, given the moderate sizes of many loci. To study how the human gp91-phox gene is expressed specifically in myeloid hematopoietic cells, we introduced yeast artificial chromosome (YAC) clones and derivatives generated in yeast into mouse embryonic stem cells competent to differentiate to myeloid cells in vitro or into mouse chimeras. Fully appropriate regulation was recapitulated with a 130-kb YAC containing 60 and 30 kb of 5′ and 3′ flanking sequences, respectively. Immunodetection of human gp91-phox protein revealed uniform expression in individual myeloid cells. The removal of upstream sequences led to decreased overall expression which reflected largely a variegated pattern of expression, such that cells were either "on" or "off," rather than pancellular loss of expression. The proportion of clones displaying marked variegation increased with progressive deletion. DNase I mapping of chromatin identified two hypersensitive clusters, consistent with the presence of multiple regulatory elements. Our findings point to cooperative interactions of complex regulatory elements and suggest that the presence of an incomplete set of elements reduces the probability that an open chromatin domain (or active transcriptional complex) may form or be maintained in the face of repressive influences of neighboring chromatin.**

The packaging of DNA into individual nucleosomes and higher-order chromatin fibers has been recognized for some time, yet the role chromatin plays in regulating gene transcription is just becoming amenable to study. Current models propose that the effects of chromatin on gene expression are not simply repressive in character (for a review, see reference 45). Instead, chromatin is a dynamic structure, undergoing constant remodeling during transcription (for a review, see reference 15). The organization of chromatin into independent transcriptional units or gene domains may aid in localizing the structural changes associated with transcriptional activity (for a review, see reference 11). Within these domains, the interplay of DNA sequences, transcription factors, histones, and other chromatin-associated proteins determines whether an active chromatin state is established and maintained. In support of these concepts, a number of regulatory elements with distinct functional properties have been identified to lie within regions of altered chromatin configuration (for a review, see reference 14). These include locus control regions (LCRs) (1, 3, 16, 26, 30, 39), enhancers (22, 29), insulators (7, 25, 43), and matrix attachment regions (8, 19, 23, 40, 64).

The dissection of *cis*-regulatory elements has provided clues to molecular mechanisms that govern transcriptional regulation. LCRs, first identified in the β-globin gene cluster (30) and now known to be associated with a limited number of other gene loci, including the lysozyme (3), CD2 (16), and visual pigment (71) genes, are thought to act by establishing and maintaining open chromatin domains. Their presence in *cis* to promoter-reporter gene constructs confers high-level, integration-site-independent, copy-number-dependent expression in transgenic mice. LCR subdomains contain arrays of transcription factor binding sites, and evidence suggests that these DNA binding sequences have distinct and separable functions. Recent studies have described at least three independent activities, one essential for facilitating factor accessibility in chromatin (13, 37), another necessary for stimulating protein-protein interactions between enhancer- and promoter-bound factors (56), and a third needed for defining domain boundaries (for reviews, see references 9 and 24). For complete transcriptional activity, LCRs appear to require all of these functional components and possibly others yet to be defined. The generation of an active transcriptional state may be viewed as a multistep process that involves competitions between "chromatin opening" proteins and nucleosomes, as well as cooperative interactions of several regulatory elements and the many transcription factors binding to them.

The organizational complexity of LCRs may reflect the requirements for stringent control of gene expression. Some elements are expected to have uniquely specified functions, while others may be redundant. A better understanding of chromatin domains necessitates further identification of *cis*-regulatory elements comprising LCRs. The paucity of LCR-like elements among characterized genes may reflect a bias imposed by inefficient detection assays. Since elements may function at great distances (>50 kb) either upstream or downstream of the proximal promoter, the identification of the entire repertoire of a gene's regulatory regions through traditional promoter-reporter approaches or even transgenic anal-

* Corresponding author. Phone: (617) 355-7910. Fax: (617) 355-7262. E-mail: ORKIN@rascal.med.harvard.edu.

† Present address: Division of Pediatric Hematology/Oncology, Stanford University School of Medicine, Palo Alto, Calif.

TABLE 1. Primer sets used in this study

| Primer set | Sequences | Location |
|---|---|---|
| Left arm | ggtgatgtcggcgatataggcgccagcaac and tccgtaatcttgagatcgggcgt | Vector end |
| Right arm | gctacttggagccactatcgactacgcgat and gtgataaattaaagtcttgcgccttaaacc | Vector end |
| A (N terminus) | gtgactggatcattatag and cccatcgatgcggccgcgtggcagaggttgaat | First exon |
| B | aaaaacaaaatgaaaccaga and aaatgcccacaaatgaataa | −16 kb |
| C | ggaaggttttgtaggggatg and gggtggggaggtatttgg | −29 kb |
| D | gaagcaagcagaggacagg and agcaaagtaggtgtctggg | −40 kb |
| 3′ UT | cagccaccacgagccagctgccagcctagc and gctcattcattttaatagttgagca | 3′ untranslated region |

ysis may be unproductive. As an alternative approach, one may consider unbiased or "top-down" strategies to *cis* element dissection in which large genomic fragments are utilized at the outset in an effort to recapitulate correct developmental and quantitative gene expression. Such approaches are based on the demonstration that the stable transfer of yeast artificial chromosomes (YACs) containing large contiguous genomic DNA fragments into mammalian cells leads to proper transgene expression (36, 61, 65). In principle, the combined use of manipulation of YACs in yeast and transfer of YACs into embryonic stem (ES) cells provides a powerful strategy with which to investigate and identify novel distant regulatory domains within a gene locus.

Here, we explore the feasibility of this approach in defining regulatory domains that govern the transcription of a moderate-sized human gene, gp91-phox. This gene, which encodes the heavy chain of a heterodimeric cytochrome, is transcribed exclusively in terminally differentiated, phagocytic hematopoietic cells (60). Given its restricted pattern of expression, it serves as a model gene with which to analyze cell regulation in myeloid blood lineages. Prior, conventional approaches to analyze gp91-phox gene expression, including the introduction of promoter-reporter gene constructs into tissue culture cell lines and the generation of transgenic mice, failed to identify regulatory elements capable of directing high-level, tissue-specific gene expression (62). These results suggested that the sequences required for appropriate transcriptional regulation reside at a distance from this gene or are dispersed.

Using YACs containing the human gp91-phox gene transferred into ES cells, we have recapitulated high-level, developmentally appropriate gene expression. As such, we have defined the domain boundaries of its autonomous transcriptional unit to 130 kb of genomic DNA contained within a single YAC (designated Y1). Deletional analysis using YACs modified in yeast and DNase I hypersensitivity mapping of myeloid cell chromatin pointed to the existence of multiple *cis* elements. Rather than leading to complete (or partial) loss of gene expression uniformly in all cells, the removal of distant sequences resulted in marked cellular heterogeneity (or variegation) of expression so that transgenes are expressed in an apparently all-or-none manner. High-level, pancellular expression was achieved only in the presence of a full set of *cis*-regulatory elements on a contiguous DNA fragment.

## MATERIALS AND METHODS

**Retrofitting vectors.** Neomycin retrofitting vectors were constructed to target the *ura3* site located on the YAC right arm for the integration of yeast *lys2* and the neomycin resistance gene. The 5.0-kb *Hin*dIII/*Eco*RI fragment of yeast *lys2* was ligated into plasmid pBSKS to yield pBS/*lys2*. PGK-neo or EF1α-neo cassettes were then added. The PGK-neo cassette was excised from plasmid pPNT (66) with *Not*I and *Eco*RI and ligated into pBS/*lys2*. The EF1α-neo cassette was excised from plasmid pBS/EF1α-neo (see below) with *Hin*dIII and *Sma*I, blunt ended with Klenow fragment, and ligated to *Sma*I-digested pBS/*lys2*. Then PGK-neo/*lys2* and EF1α-neo/*lys2* cassettes were excised with *Hin*dIII/*Not*I and *Hin*dIII/*Sma*I, respectively, blunt ended, and ligated to the blunted *Stu*I site in plasmid B790 (*ura3*-containing pUC18-derived vector). Subsequent digestion with *Hin*dIII released the PGK-neo/*lys2* and EF1α-neo/*lys2* cassettes flanked by regions of the yeast *ura3* gene.

The hygromycin retrofitting vector was constructed as follows. The *lys2* gene was excised from pBS/*lys2* with *Sal*I and *Sma*I, blunt ended with Klenow fragment, and ligated into pRMM352 (a vector containing the PGK-hygro cassette) (50). The resulting plasmid was digested with *Hin*dIII and *Cla*I to obtain a fragment containing both selectable markers. The fragment was then subcloned into alkaline phosphatase-treated, *Stu*I-digested B790.

EF1α-neo was constructed as follows. The Neo$^r$ gene cassette was generated by PCR with primers ttgcggccgccaatatgggatcggcc and ttgcggccgcgaattcgagctcggt acc. Both primers were synthesized to include *Not*I sites to facilitate cloning of the PCR product. The *Not*I-digested PCR product was ligated into pEF-BOS, a mammalian expression vector which utilizes the promoter from the human EF-1α chromosomal gene (48). The resulting plasmid was digested with *Hin*dIII, blunted with Klenow fragment, digested with *Sma*I, and ligated to either pICU or pLUS (two yeast integrating vectors that target yeast sequences located on the left and right arms of YACs, respectively) (31) or inserted into the *Sma*I site of pBS/*lys2*, as described above.

**"Fragmenting" vectors.** Fragmenting vectors were modifications of pBP103 (53) and included a yeast telomeric sequence, a yeast selectable marker, a neomycin cassette, and a polylinker into which several target elements were added. Neomycin cassettes, PGK-*neo* and EF1α-*neo*, were excised from pPNT and pBS/EF1α-neo with *Not*I/*Xba*I and *Hin*dIII/*Sma*I, respectively, and inserted at the *Not*I/*Xba*I and blunted *Xba*I sites of pBP103, respectively. pBP108 and pBP109 contain human *Alu* repeat elements (a 0.3-kb *Alu* sequence from Blur8) inserted at the *Bam*HI site of pBP103 in both orientations. Additional target elements inserted into the polylinker of pBP103 included a 2-kb *Eco*RI/*Xba*I fragment and a 2-kb *Bam*HI fragment derived from gp91-phox upstream flanking sequences at approximately −14 and −29 kb, respectively. All constructs were linearized at *Sal*I to facilitate homologous recombination.

**PCR primers.** Multiple primer sets were used to demonstrate the presence or absence of specific regions spanning the human gp91-phox gene locus in YAC derivatives as well as YAC-integrated ES clones. The designations and descriptions of these primers are given in Table 1.

**Yeast strains and manipulations.** Two *Saccharomyces cerevisiae* AB1380 strains carrying a 130-kb Y1 or 80-kb Y3 YAC containing the human gp91-phox gene were gifts of Tony Monaco. Strains were propagated in synthetic minimal media without tryptophan and uracil (2). Y1- and Y3-containing yeast strains were transformed with retrofitting vectors to incorporate the Neo$^r$ gene by a spheroplast protocol (6), and transformants were selected on minimally supplemented plates lacking tryptophan and lysine. Retrofitted YACs were subsequently transferred by electroporation into yeast strain YPH857 (*MAT*α *ade2-101 cyh2*$^r$ *leu2Δ1 lys2-801 his3Δ200 trp1Δ63 ura3-52*), a host strain suitable for YAC fragmentation. YAC strains were transformed by the spheroplast method with the fragmenting vectors described above and selected on plates lacking tryptophan and histidine. The sizes of YACs were determined by pulsed-field gel electrophoresis and Southern blot analysis with probes spanning the human gp91-phox gene locus. An approximate location of deletion endpoints was achieved by PCR with the primer sets described above, including sequences from the first exon, distal upstream sites at −16, −29, and −40 kb, and the 3′ untranslated region.

**Cell culture.** The mouse ES cell line CCE (57) was cultured on gelatin-coated (0.1%) dishes and grown in high-glucose Dulbecco modified Eagle medium (DMEM) with 15% fetal calf serum (FCS) at 37°C with 5% $CO_2$. Conditioned medium was supplemented with ESgro (10,000 U/ml), 2 mM glutamine, 0.1 mM nonessential amino acids, and penicillin-streptomycin (50 U/ml). Human promyelocytic HL60 cells were grown in RPMI 1640 medium containing 10% FCS, 2 mM glutamine, and penicillin-streptomycin.

**Spheroplast fusion.** Fusions were performed by the method of Huxley et al. (34). YAC-containing yeast cells were inoculated in 50 ml of selection medium and incubated at 30°C to an optical density at 600 nm of 0.5 to 1.0. Cells were collected by centrifugation and washed once in sterile water and once in 1 M sorbitol. Cells resuspended in 20 ml of SCE (1 M sorbitol, 0.1 M sodium citrate [pH 5.8], 10 mM EDTA [pH 8.0])–43 μl of β-mercaptoethanol–0.4 mg of Zymolase 100T were incubated at 30°C until 90% of yeast cells had formed spheroplasts. Spheroplasts were pelleted by centrifugation, washed twice in STC (1 M

sorbitol, 10 mM Tris base [pH 7.5], 10 mM $CaCl_2$), and resuspended in a final volume of 2 ml of STC.

Exponentially growing ES cells were harvested at 50 to 75% confluency, washed twice in serum-free DMEM, and layered onto a yeast spheroplast pellet containing $10^7$ cells. After centrifugation, cells were mixed by gentle tapping and incubated at room temperature in 500 μl of polyethylene glycol 1500 solution (50% polyethylene glycol in 75 mM HEPES [pH 8.0], 5 mM $CaCl_2$). After a 2-min incubation, cells were diluted in 5 ml of serum-free DMEM and incubated for an additional 30 min. Cells were collected by centrifugation, resuspended in complete DMEM medium, and distributed into four 10-cm-diameter tissue culture dishes. Twenty hours after fusion, the medium was removed and replaced with complete medium containing 200 μg of active G418 per ml.

**Electroporation.** Preparation of YAC DNA was performed by the method of Cuoto et al. (10) with only minor modifications. Yeast cultures (100 ml) were grown to high density in selective medium at 30°C. Cells were pelleted by centrifugation, washed once in 50 mM EDTA (pH 8), and embedded in preparative 1% low-melting-point agarose. YAC DNA was separated from yeast chromosomal DNA by pulsed-field gel electrophoresis with a 1% agarose gel run at 200 V with a switch time of 0.5 to 8.6 s. YAC DNA was excised from the gel, treated with β-agarase, and concentrated with a Centriprep 100 column. Twenty-five microliters of this DNA solution was used to electroporate $10^7$ ES cells in 0.8 ml of HEPES-buffered saline at 450 V and 25 μF. Cells were allowed to recover in complete medium for 24 h and then selected in G418-containing medium until colonies appeared.

**Analysis of ES clones.** Genomic DNA from G418-resistant clones was prepared with an Applied Biosystems DNA extractor and analyzed for YAC integration by Southern analysis and PCR. The presence of YAC vector arms was detected by PCR with primers specific to right and left arm sequences. Additional primer pairs corresponding to sequences from the N terminus, a distal upstream site at −29 kb, and the 3′ untranslated region were used to demonstrate the integrity of YAC sequences. Southern blots were prepared with *Bam*HI/*Eco*RV-digested genomic DNA derived from G418-resistant ES clones and hybridized successively to probes subcloned from the 5′ and 3′ untranslated regions and the first exon.

**In vitro differentiation (two step).** Cells were dissociated by trypsinization and cultured in 1% methylcellulose in IMDM containing a 15% serum mixture (1:1 mix of plasma-derived serum and FCS), glutamine, monothioglycerol, kit ligand, and interleukin-11 (IL-11). Between 1,000 and 2,500 CCE ES cells were plated onto 35-mm-diameter bacterial-grade dishes and incubated at 37°C with 5% $CO_2$. Hematopoietic precursors within embryoid bodies (EBs) were replated at various stages of development. EBs were collected by dilution of methylcellulose with IMDM, and cells within EBs were disaggregated by incubation (90 min at 37°C) in a phosphate-buffered saline (PBS) solution containing 0.25% collagenase and 20% FCS. After incubation, cells were passed through a syringe with a 20-gauge needle. EB-derived cells were counted and replated in methylcellulose cultures similar to those described earlier, except that 10% serum was used and additional growth factors were added. Erythoid cultures contained erythropoietin and transferrin; myeloid cultures contained IL-1, IL-3, granulocyte-macrophage colony-stimulating factor (GM-CSF), and granulocyte colony-stimulating factor (G-CSF).

**RNA and cDNA preparation.** Total RNA was isolated from in vitro-differentiated ES cultures with RNAzol. cDNA synthesis was performed by the addition of a reverse transcription (RT) mix [containing 100 mM KCl, 20 mM Tris-HCl (pH 8.4), 10 mM dithiothreitol, 5 mM $MgCl_2$, 1 mM (each) deoxynucleoside triphosphates, 1 U of RNasin (Promega), oligo(dT) primer, and Moloney murine leukemia virus reverse transcriptase] to the RNA solution. After an incubation of 60 min at 42°C, the samples were either subjected to PCR or stored at −20°C. PCR was performed with two sets of primers, one which detects the endogenous mouse gp91-phox transcript (gatgatgggcctaagtataa and aataccggtcagaaatccc) and another for the exogenous human transcript (gatattccacctaagttctt and ttttttgttca caaactgt), under the following conditions: 94°C for 1 min, 48°C for 2 min, and 72°C for 2 min for 30 cycles, followed by 72°C for 7 min. Ten microliters of the resulting amplified material was gel electrophoresed in 1× TAE (0.04 M Tris-acetate, 0.001 M EDTA) buffer. Semiquantitative RT-PCR was performed with reverse-transcribed RNA from differentiated ES clones by using primers (aggg gttccagtgcgtg and gggcacttgacaaaaatgta) common to human and mouse sequences. cDNA products were amplified by PCR in the presence of tracer [$\alpha$-$^{32}$P]dCTP for 20, 22, 24, and 26 cycles. PCR products were digested with *Eco*RV and separated by gel electrophoresis on 6% acrylamide gels to reveal an *Eco*RV restriction fragment-length polymorphism. The results were visualized by autoradiography and quantitated by PhosphorImager (Molecular Dynamics) analysis of the dried gel.

**Immunofluorescence.** Colonies of differentiated cells were picked with a micropipette, placed in PBS, and spun in a Shandon Cytospin 2 at 350 rpm onto cytocentrifuge slides. Cells were fixed in methanol for 1 to 2 min and allowed to dry briefly at room temperature. Fixed cells were incubated overnight at 4°C with a mouse monoclonal antibody raised against human gp91-phox, MAb48 (67), diluted 1:400 in PBS containing 5% goat serum, 0.2% Tween 20, and 0.5% Triton X-100. Cells were washed three times in PBS for 5 min before the addition of second antibody. Primary antibody binding was detected by incubating cells with 1:400 dilution of Cy3-conjugated goat anti-mouse immunoglobulin G antibody for 60 min at room temperature. Cells were subsequently washed three times with PBS for 5 min, mounted with Immumount, and visualized under a fluorescent microscope.

**DNase I hypersensitivity site mapping.** Cells in exponential-growth phase were harvested, washed twice in ice-cold PBS, and resuspended in 6 ml of ice-cold buffer I (300 mM sucrose, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 15 mM Tris-HCl [pH 7.4], 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, and 0.1 mM EDTA) at $10^7$ cells per ml. Cells were lysed by the addition of 0.12 ml of 10% Nonidet P-40, and nuclei were pelleted by low centrifugation over a sucrose cushion (buffer I plus 1.7 M sucrose). The nuclear pellet was resuspended in buffer III (buffer I plus 5% glycerol) at $2 \times 10^7$ nuclei per ml. Nuclease digestion was carried out in a volume of 100 μl. Nuclei were incubated at 26°C with either no nuclease or increasing amounts of DNase I (Boehringer Mannheim) (0, 1.25, 1.875, 2.5, 3.75, 5, 7.5, 10, and 15 U per reaction mixture) for 4 min. Digestion was stopped by the addition of EDTA to 10 mM. Nuclei were lysed with sodium dodecyl sulfate to 0.5%, followed by proteinase K (100 mg/ml) treatment at 37°C overnight. DNA was extracted with phenol-chloroform several times, precipitated with ethyl alcohol, and digested to completion with the indicated restriction enzymes. Digested DNA was electrophoresed, transferred onto Hybond $N^+$ membranes by the Southern method, and hybridized sequentially to various probes subcloned from upstream flanking sequences.

## RESULTS

**Experimental strategy.** We have established a system that uses YAC transfer into mammalian cells, manipulation of YACs in yeast, and ES cell technology to identify distal regulatory elements necessary for directing proper gp91-phox gene expression both in vitro and in vivo. YACs Y1 (130 kb) and Y3 (80 kb) contain an intact human gp91-phox locus with approximately 60 and 20 kb of 5′ sequences and 30 and 20 kb of 3′ sequences, respectively (Fig. 1A). To facilitate selection of YAC-integrated ES clones, both YACs were retrofitted to include a neomycin resistance gene under the control of either the PGK or EF1-α promoter. The neomycin resistance gene, together with the yeast *lys2* gene, was introduced by site-specific recombination into the yeast *ura3* gene located on the YAC right arm. Properly retrofitted YACs were selected by their ability to grow on selective medium (i.e., $Trp^-$ $Lys^-$) and subsequently transferred into mouse ES cells by spheroplast fusion and electroporation (34) as described in Materials and Methods.

To examine whether the human gp91-phox locus from YACs Y1 and Y3 was transferred in its entirety, PCR and Southern blot analyses were performed. Genomic DNA derived from YAC-integrated ES clones was prepared and subjected to PCR analysis with primer sets specific to vector sequences from the left and right YAC arms. Amplification of yeast sequences in the transfected lines indicated the integration of intact YAC ends (Fig. 1B). Furthermore, internal sequences spanning regions of the YAC were also examined by PCR with primers corresponding to the N terminus, a distal upstream site, and the 3′ untranslated region (Fig. 1C). As expected, amplification of these sequences was observed in the transfected ES lines, with the exception of the distal upstream site in Y3 clones, a result consistent with the fact that these distal 5′ sequences are not present in YAC Y3. Southern blot analyses diagnostic for fragments spanning the gp91-phox locus were also performed. Filters were probed with radioactively labeled DNA fragments from upstream, downstream, and coding sequences (Fig. 1D). All clones obtained by spheroplast fusion (denoted with a numerical suffix, e.g., Y1N.1) were found to have intact YAC DNA. Clones derived by electroporation (designated by an alphabetical suffix, e.g., Y1N.A) also possessed structurally intact YACs, with the only exception of Y1N.B. In summary, detailed PCR and Southern analyses demonstrated that the vast majority of ES clones contained single-copy (data not shown), unrearranged YAC transgenes.

Cells of multiple hematopoietic lineages are obtained upon in vitro differentiation of ES cells (41, 73). Using a two-step in vitro differentiation protocol, we investigated lineage specific-

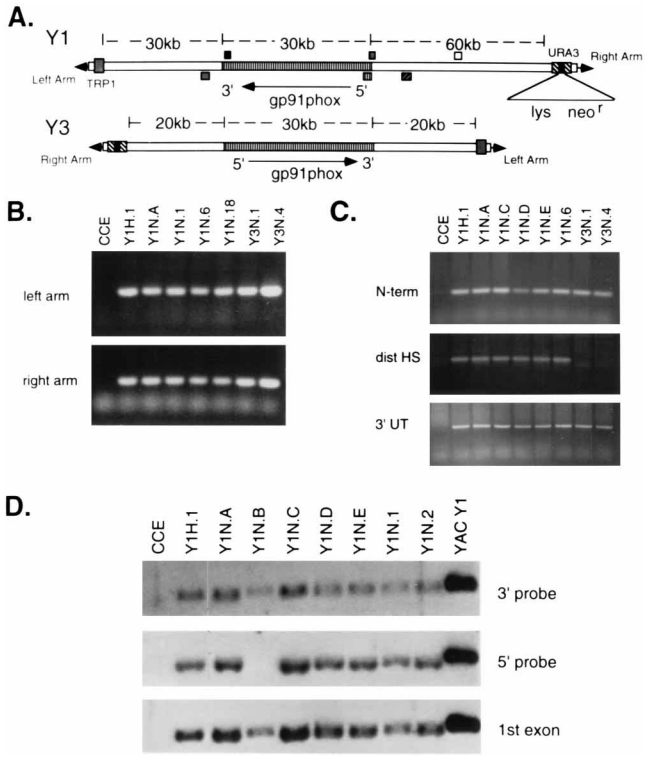



FIG. 1. Characterization of YAC-integrated ES clones. (A) Schematic representations of gp91-phox YACs Y1 and Y3 retrofitted with Neo$^r$ gene. Hatched boxes represent approximate locations of regions amplified by PCR (above YAC Y1) and various probes used in Southern analysis (below YAC Y1). (B) PCR analysis of wild-type ES line CCE and YAC-containing ES clones with primers specific to left and right arm YAC sequences. (C) PCR amplification of YAC-transfected ES clones corresponding to sequences from the N terminus (N-term), a distal upstream site at −29 kb (dist HS), and the 3′ untranslated region (3′ UT). (D) Representative Southern blots of *Bam*HI- and *Eco*RV-digested genomic DNA derived from wild-type ES line CCE, YAC-integrated ES clones, and YAC Y1. Filters were sequentially hybridized with probes subcloned from the 5′ and 3′ untranslated regions and the first exon.

ity and transgene expression levels among YAC-integrated ES clones. ES cells were cultured in methylcellulose containing kit ligand and IL-11 to generate EBs that harbor hematopoietic precursors. These EBs were disaggregated with cellulase and replated in methylcellulose cultures containing cytokines necessary for the formation of terminally differentiated colonies. Cytokines included erythropoietin plus kit ligand for erythroid colonies and IL-1, IL-3, GM-CSF, and G-CSF for myeloid (macrophage and neutrophil) colonies. Erythrocyte, macrophage, and neutrophil colonies developed efficiently in ES clones derived from both spheroplast fusion and electroporation.

**YAC Y1, but not Y3, behaves as an independent regulatory unit.** After in vitro differentiation of YAC-integrated ES clones, human gp91-phox expression was examined at both RNA and protein levels. Exogenous (human gp91-phox) and endogenous (mouse gp91-phox) transcripts were detected by RT-PCR. RT-PCR analysis was performed with RNA extracted from pools of in vitro-differentiated ES colonies by using primer sets unique to the human and mouse transcripts. Human gp91-phox transcripts were detected in all ES lines containing intact YACs and, as expected, were absent in wild-type ES cells and undifferentiated YAC integrants (Fig. 2). As anticipated, endogenous mouse gp91-phox transcripts were detected only in differentiated ES cells. These initial findings demonstrate that gp91-phox YACs integrated into ES cell genomes are expressed in an appropriate developmentally regulated pattern.

Semiquantitative RT-PCR was used to examine the levels of transcripts among Y1 and Y3 clones. Reverse-transcribed RNA was subjected to PCR with primers common to both mouse and human gp91-phox sequences. PCR reactions were performed in the presence of tracer [$^{32}$P]dCTP to permit quantitation. Mouse and human PCR products were distinguished by *Eco*RV digestion, revealing two mouse subfragments and one human product (Fig. 3). The ratio of human to mouse gene products reflects the relative levels of expression from each clone. Y1 clones were expressed consistently at high levels, equivalent to or exceeding that of the endogenous mouse gene, whereas Y3 derivatives were expressed at much lower levels. A comparison of ratios among Y1 and Y3 clones indi-

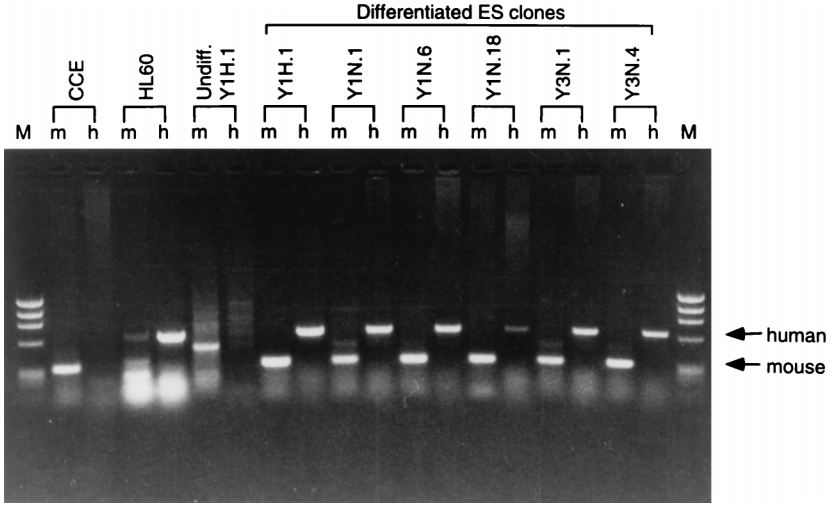
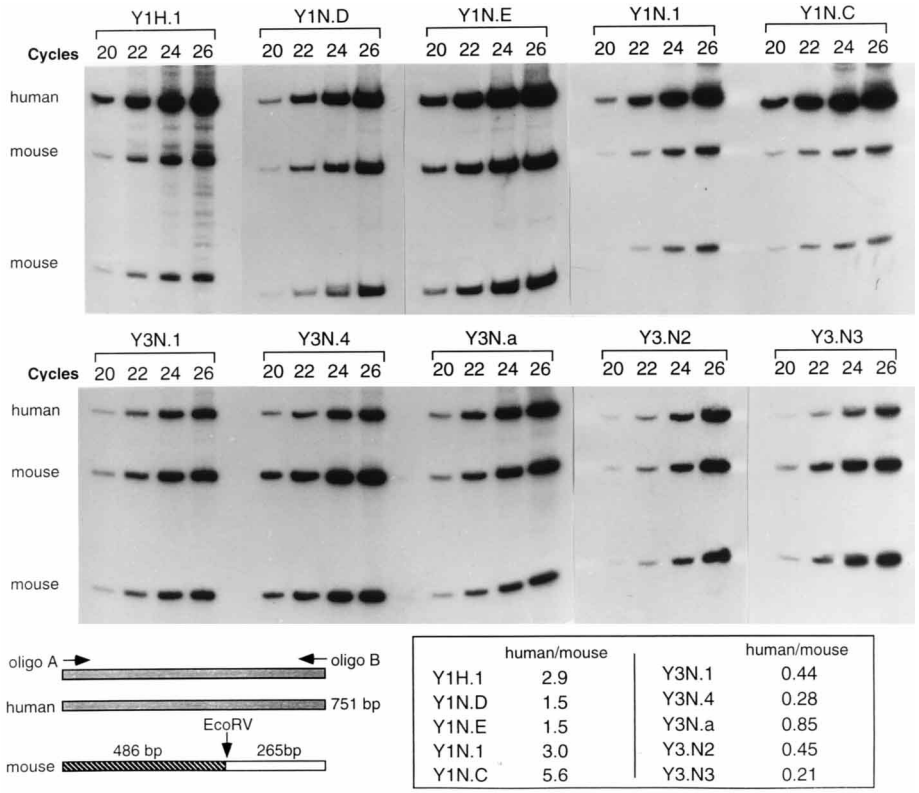



FIG. 2. Expression of the human and mouse gp91-phox genes in YAC-containing ES lines differentiated in vitro. RNAs harvested from differentiated ES colonies were reverse transcribed and amplified with primers specific to human (h) and mouse (m) sequences, yielding products of 700 and 400 bp, respectively. As controls, RNAs prepared from wild-type ES line CCE, human promyelocytic line HL60, and undifferentiated (Undiff.) YAC-integrated ES clone Y1H.1 were also analyzed. Lanes M, markers.




| | human/mouse | | human/mouse |
|---|---|---|---|
| Y1H.1 | 2.9 | Y3N.1 | 0.44 |
| Y1N.D | 1.5 | Y3N.4 | 0.28 |
| Y1N.E | 1.5 | Y3N.a | 0.85 |
| Y1N.1 | 3.0 | Y3.N2 | 0.45 |
| Y1N.C | 5.6 | Y3.N3 | 0.21 |

FIG. 3. Semiquantitative analysis of human gp91-phox gene expression among ES clones containing YACs Y1 and Y3. RNAs were prepared from differentiated ES lines, reverse transcribed, and amplified by PCR with primers common to human and mouse transcripts. PCR products were removed after 20, 22, 24, and 26 cycles and digested with *Eco*RV to reveal one human and two mouse products. Quantitation was performed by PhosphorImager analysis, and data are ratios between the human and mouse products. oligo, oligonucleotide.

cated that Y1 integrants were expressed at approximately five-fold-greater levels (Fig. 3). The higher levels of transcription in Y1 clones suggests that Y3 lacks some distal regulatory elements that contribute to full expression.

To determine more accurately whether expression of the YAC transgene was specific to cells of myeloid lineage, immunofluorescence assays were performed with in vitro-differentiated ES colonies and a monoclonal antibody (MAb48) (67) generated against the human gp91-phox protein. Neutrophil, macrophage, and erythrocyte colonies were selected and centrifuged onto microscope slides for May-Grunwald/Giemsa staining and immunofluorescence assays. Immunostaining was specific to neutrophils and macrophages (Fig. 4A); no staining was observed in erythrocytes, mirroring the observed pattern for endogenous human gp91-phox expression. Consistently greater than 90% of myeloid cells stained for human gp91-phox protein among Y1 clones; however, the percentage of cells staining among Y3 clones varied. These results further suggest that YAC Y1, but not Y3, contains the full complement of regulatory elements required for high-level, tissue-specific expression.

Although the in vitro ES differentiation system closely parallels mouse hematopoietic development (41), we sought to verify that the Y1 YAC transgene is functional and appropriately regulated in vivo. To address this goal, Y1-integrated ES clones were injected into blastocytes and levels of chimerism of 10 to 30%, as judged by coat color, were generated in four mice. The presence of YAC DNA was confirmed by PCR and Southern blot analyses of tail DNA. Peripheral blood samples derived from these mice contained myeloid cells expressing the human gp91-phox protein, as detected by immunofluorescence (Fig. 4A). Furthermore, bone marrow cells were harvested and cultured in methylcellulose containing kit ligand, IL-1, IL-3, GM-CSF, and G-CSF. By using procedures similar to those described above, transgene expression was detected in differentiated colonies by RT-PCR and found to be present only in myelomonocytic cells by immunofluorescence (not shown).

In summary, ES clones containing YAC Y1, but not Y3, exhibited proper human gp91-phox gene regulation. High-level, tissue-specific expression was observed among Y1 clones both upon in vitro differentiation of ES cells and within chimeric mice. Hence, the expression of this gene is recapitulated in vivo after introduction of a genomic fragment of ~130 kb.

**A yeast genetic approach allows rapid construction of YAC deletion derivatives.** High-frequency recombination in *S. cerevisiae* facilitates the manipulation of YAC DNA (for a review, see reference 4). To assess the effects of specific deletions within the gp91-phox locus on expression, we performed deletion analysis of flanking sequences contained on YAC Y1 by homologous integrative transformation (52, 68). A detailed set of 5′-terminal deletions were generated by targeting the human *Alu* repeat elements distributed throughout the YAC for homologous recombination with a fragmenting vector containing yeast and mammalian selectable markers and yeast telomeric sequences (53). The integration of a telomere at the targeted *Alu* site resulted in the loss of YAC sequences distal to the site of recombination (Fig. 5A). Since *Alu* elements are found on average approximately every 5 to 10 kb (51), deletion derivatives of various sizes can be obtained. In addition, by replacing the *Alu* element with unique target sequences, we

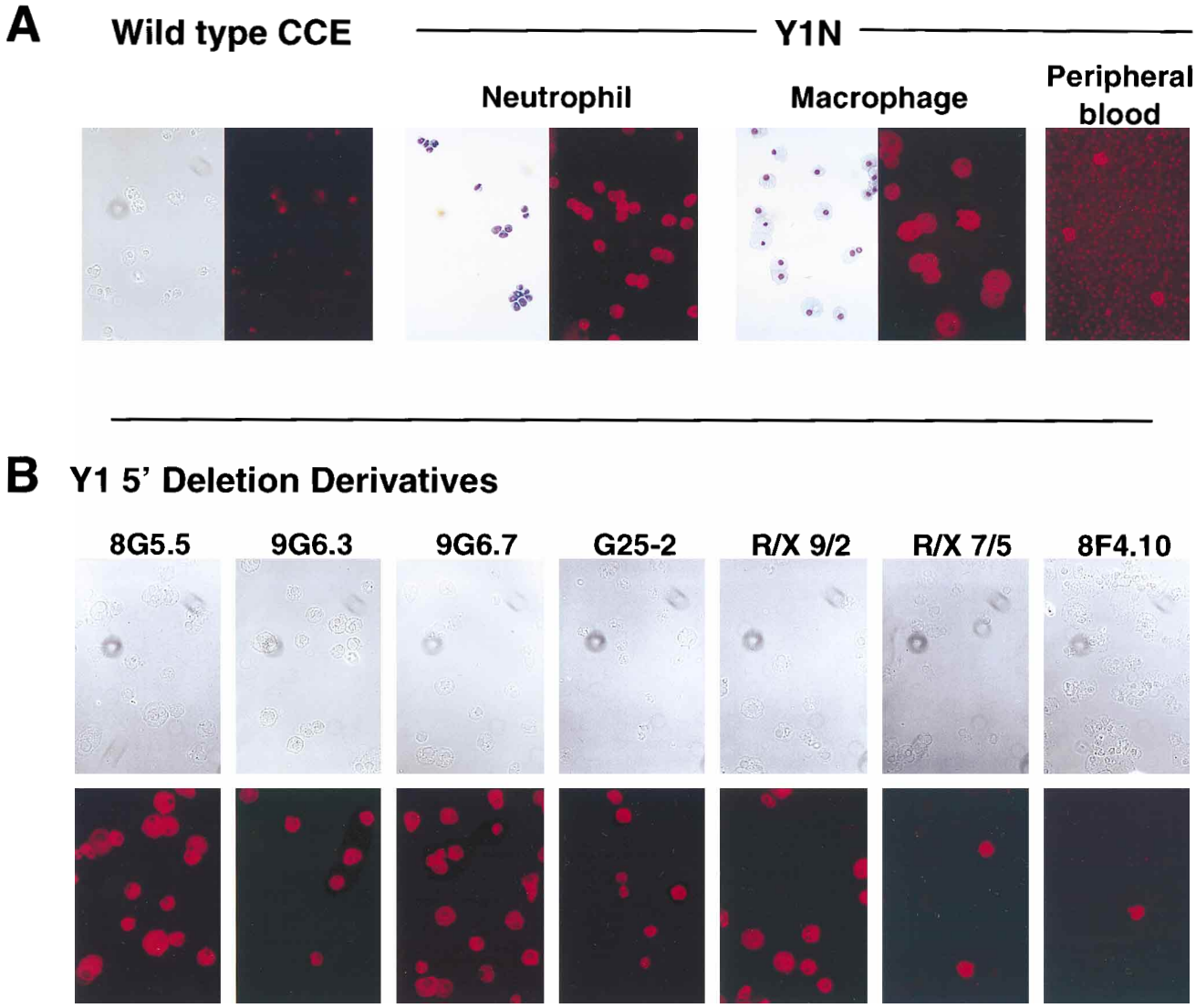



FIG. 4. Immunofluorescence of individual YAC-integrated ES clones in vitro differentiated along a myeloid pathway. Transgene expression was detected with a monoclonal antibody (MAb48) raised against the human gp91-phox protein. (A) Representative analysis of a Y1 YAC clone reveals no staining among myeloid cells derived from untransfected ES cells (left), uniform, high-level staining in neutrophils and macrophage from in vitro-differentiated clone Y1N (middle), and staining specific to myeloid cells in a peripheral blood smear of a chimeric mouse containing Y1N DNA (right). Note that panels adjacent to fluorescent images depict either cells from identical fields seen under phase-contrast microscopy or cells stained with May-Grunwald/Giemsa stain for morphological assessment. (Note also that the punctate staining displayed in the peripheral blood smear is nonspecific staining of nucleii.) (B) A subset of 5′ deletion derivatives analyzed by immunofluorescence. All clones displayed a heterocellular (variegated) pattern of expression; however, clones with progressively larger deletions expressed the transgene in much lower percentages of cells (left to right). Top and bottom rows represent identical fields seen under phase-contrast and fluorescent microscopy, respectively, and aid in determining the number of cells stained.

were able to fragment YAC Y1 at specific sites where *Alu* sequences were lacking (Fig. 5A). Pulsed-field gel electrophoresis and Southern blot analyses identified YAC derivatives ranging from 70 to 110 kb (Fig. 5B). An approximate determination of the deletion endpoints from these derivatives was achieved through the use of various upstream probes (not shown) and numerous PCR primer sets (Fig. 5C). These PCR primers encompassed regions at the first exon, several upstream sequences located at approximately −16, −29, and −40 kb, and the 3′ untranslated region. Taken together, our PCR analyses suggest that these YAC derivatives varied with respect to the amount of 5′ flanking sequences, with some clones containing as little as 4 kb (8F5 and 8F4) and others possessing as much as 40 kb (8G1). In addition, R/X and G25, derivatives obtained by targeted fragmentation with unique genomic fragments subcloned from −15- and −29-kb upstream regions, served as useful size markers and further aided in determining the positions of upstream deletions.

**5′ deletion derivatives display variegation of expression and loss of position independence.** To identify upstream regulatory sequences, we analyzed YAC derivatives for expression in ES cells differentiated in vitro. YAC derivatives were introduced stably into ES cells by spheroplast fusion, a procedure which results in efficient transfer of intact and unrearranged YAC DNA (see above). Expression among 5′ deletion derivatives was monitored at the single-cell level by immunofluorescence (see Materials and Methods) and quantitated by determining the number of stained cells in a total population of 100 cells.

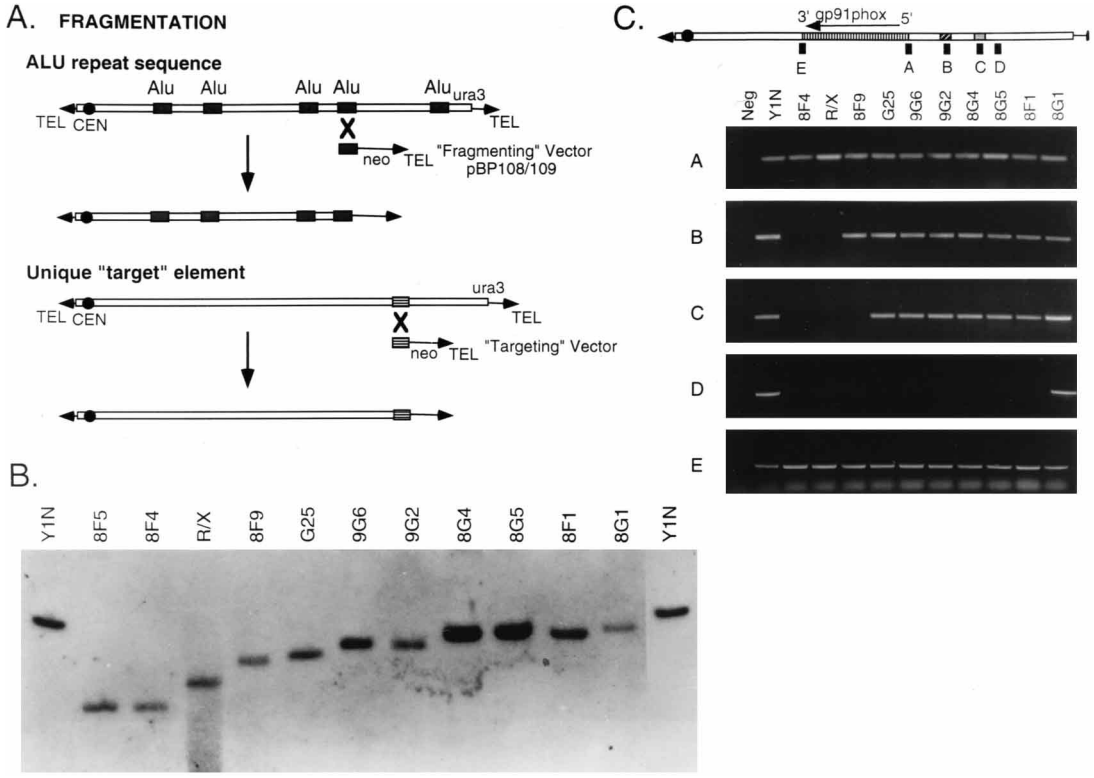



FIG. 5. Generation of deletions of YAC Y1 by targeted homologous recombination in *S. cerevisiae*. (A) Schematic representation and strategy to generate deletion derivatives by YAC fragmentation. Recombination between *Alu* repeats on fragmenting vectors pBP108 and pBP109 and corresponding sequences on Y1 results in the introduction of a telomere (TEL) and subsequent loss of all sequences distal to the site of recombination. The presence of multiple *Alu* repeats allows for many potential sites of recombination. Similarly, specific regions on Y1 may be targeted for replacement by using uniquely specified sequences as target elements, as depicted with a targeting vector. Proper homologous recombination leads to the removal of sequences distal to the target element. (B) Set of deletion derivatives generated by targeted transformation of Y1 with pBP108- and pBP109-derived fragmenting and unique targeting vectors. YAC derivatives were separated by pulsed-field gel electrophoresis, and the blot was hybridized to a fragment from the neomycin gene. Y1 is the parental YAC. All others are deletion derivatives ranging from 70 to 110 kb. (C) PCR analysis of YAC derivatives with primers corresponding to sequences from the first exon (A), upstream sites at −16 (B), −29 (C), and −40 (D) kb, and the 3′ untranslated region (E). Neg, negative control.

As noted above, Y1-integrated ES clones demonstrated high-level, tissue-specific, position-independent gp91-phox gene expression, thereby establishing the minimal set of control elements sufficient for directing appropriate transgene expression. Among these Y1 clones, 90 to 100% of macrophage and granulocytic cells expressed gp91-phox protein (Fig. 4A). The removal of 5′ flanking sequences resulted in striking effects. Large variations in transgene expression were observed not only among different YAC derivatives (8F4 versus 8G5) but also among individual clones derived from the same 5′ deletion derivative (R/X 7/5 versus R/X 9/2) (Fig. 4B). While loss of position independence was exhibited among all 5′ deletion clones analyzed, derivatives with progressively smaller flanking sequences exhibited lower levels of transgene expression on average, suggesting that expression is affected not only by chromosomal position of YAC integration but also by the loss of multiple *cis*-regulatory elements.

As shown in Fig. 4B, deletions far upstream (~−30 to −40 kb [Fig. 6]) led to a heterocellular pattern of expression, in which the majority of cells in the majority of clones expressed. This consistent, high-level expression evident by immunofluorescence was confirmed by semiquantitative RT-PCR assay (not shown). We suggest that clones in this group lack elements conferring position independence yet still retain enhancer-like sequences capable of creating and maintaining a chromatin domain permissive for transcriptional activity.

Variegated expression, although seen among all 5′ deletions, became more apparent in clones with larger deletions (Fig. 4B and 6). For example, most ES clones derived from YACs containing 30 to 40 kb of 5′ sequences (such as 8G4, 8G5, 8F1, and 8G1) expressed at levels only slightly lower than those derived from Y1 (Fig. 4B and 6). Clones with larger deletions displayed greater variegation. The majority of deletion derivatives with 15 to 30 kb of 5′ flanking sequences present (such as those derived from YACs G25, 8F9, R/X, and Y3) expressed the transgene in a smaller fraction of cells (Fig. 4B and 6). In general, YAC clones of this size were unable to overcome the silencing effects associated with inactive chromatin most likely due to the loss of multiple regulatory elements. As further evidence that the overall level of expression was determined by the number of expressing cells and not by the level of activity within expressing cells, we noted that clones with <4 kb of 5′ sequences (such as 8F4) exhibited expression in <1% of its phagocytes (Fig. 4B and 6).

**Presence of multiple hypersensitive sites suggests multiple upstream regulatory elements.** DNase I-hypersensitive-site mapping studies were performed to complement the deletion analyses and delineate potential regions containing critical *cis* elements. Nuclei prepared from promyelocytic HL60 and non-hematopoietic HeLa cell lines were treated with increasing concentrations of DNase I. Genomic DNAs prepared from these nuclei were digested with various restriction enzymes,

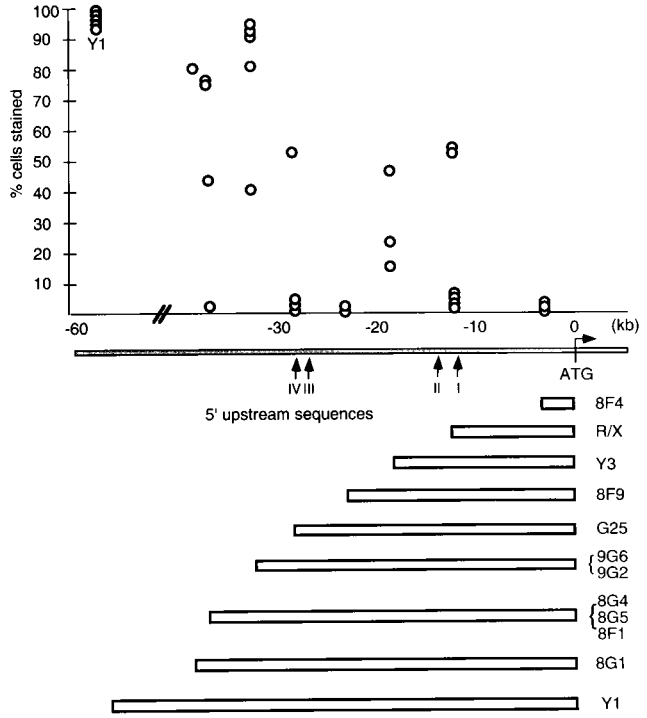



FIG. 6. Summary of 5′ deletion clones analyzed by immunofluorescence. The expression of each YAC deletion derivative is plotted as the percentage of myeloid cells staining positively against the amount of 5′ flanking sequences present. Percentages were determined by counting the number of stained cells in a total population of ~100 cells from 10 different fields. Each symbol represents a separate clonal line derived either from different YAC derivatives (i.e., different position on the *x* axis) or from the same YAC derivative (i.e., same position on the *x* axis). YAC derivatives 9G6 and 9G2, as well as 8G4, 8G5, and 8F1, have been grouped together since the amounts of 5′ sequences among these clones are similar. Notice the consistent, high-level expression exhibited among Y1 clones and deletion derivatives with >30 kb of 5′ flanking sequences. Notice also that clones with smaller flanking sequences displayed greater variegation of expression.

fractionated by gel electrophoresis, transferred to nylon membranes, and hybridized with unique, upstream DNA fragments. Probe 11C N/R hybridizes to the 5′ end of a 5-kb *Eco*RI fragment which maps to a region from 12 to 17 kb upstream of the gp91-phox gene. In HL60 nuclei, progressive nuclease digestion led to a reduction in the intensity of the 5 kb-band and the subsequent appearance of two new subbands of about 4.3 and 2.3 kb (Fig. 7A). These hypersensitive sites, designated HS I and HS II, were positioned at −13 and −15 kb, respectively. An examination of nuclease hypersensitivity further upstream of HS II revealed two additional strong hypersensitive sites. Blots prepared with *Eco*RV-digested DNA were probed with 17 X/N B, which hybridizes to a 15-kb *Eco*RV fragment corresponding to the region from −17 to −32 kb. Nuclease treatment revealed two subbands of ~7 to 8 kb, indicative of DNase I hypersensitivity at approximately −28 (HS III) and −29 (HS IV) kb (Fig. 7B). DNA digested with *Hin*dIII and hybridized with 17 X/N B confirmed the presence of HS III and HS IV at those positions (Fig. 7C). With probes upstream of HS IV, no additional sites of DNase I hypersensitivity were observed, and at ~−50 kb, general DNase I sensitivity was greatly diminished. Furthermore, analysis of DNA from HeLa nuclei failed to detect the presence of HS I, HS II, HS III, and HS IV. Our results demonstrate the presence of four tissue-specific hypersensitive sites residing in the upstream sequences of the gp91-

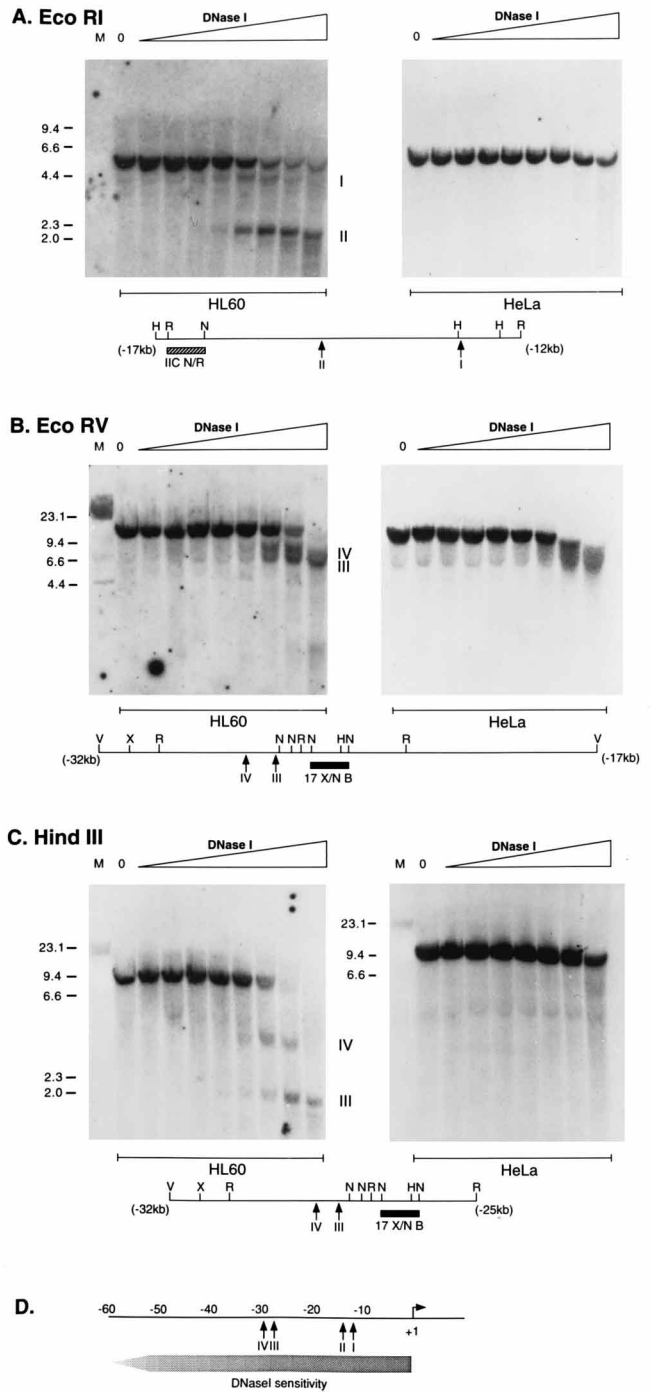



FIG. 7. DNase I-hypersensitive-site mapping of human gp91-phox upstream sequences. (A through C) Examples of autoradiographs used to identify hypersensitive sites in nuclei isolated from HL60 and HeLa cells, as described in Materials and Methods. DNA was extracted and digested with *Eco*RI, *Eco*RV, and *Hin*dIII, separated by gel electrophoresis, and transferred onto Hybond $N^+$ membranes. Blots were hybridized with DNA fragments subcloned from gp91-phox upstream sequences. A diagram of the corresponding region is shown below each blot, with the location of the probe (hatched boxes), hypersensitive sites (arrows), and relevant restriction enzyme sites (H, *Hin*dIII; R, *Eco*RI; N, *Nco*I; V, *Eco*RV; X, *Xba*I). (D) Locations of hypersensitive sites within the gp91-phox 5′ flanking sequences.

phox gene (Fig. 7D). Taken together with the results from our deletion studies, these data provide evidence for the presence of multiple regulatory domains and suggest that the establishment and maintenance of an active transcriptional complex constitute a multistep process.

## DISCUSSION

An active transcriptional unit consists of multiple *cis*-regulatory elements, often distributed over a considerable region of DNA. Current models suggest that establishment of active transcription occurs through a multistep process, in which chromatin domains are reorganized, specific factors are recruited, and protein-protein interactions are established. A number of distal regulatory sequences, such as LCRs, enhancers, matrix attachments regions, and insulators have previously been identified (14). Their functions in altering chromatin and interacting with promoter elements have often been studied through minigene constructs in transfected cell lines and/or transgenic mice. The use of large DNA constructs, such as YACs, constitutes an alternative approach to the study of gene regulation. The large cloning capacity of YACs and the efficient means for manipulating YACs in yeast offer distinct advantages over conventional minigene constructs. Here we employed this strategy to dissect elements within the human gp91-phox locus required for high-level, developmentally appropriate expression.

**Establishment of YAC-based assay system.** Our results demonstrate that a 130-kb YAC covering 40 kb of the human gp91-phox gene, 60 kb of 5′ flanking sequences, and 30 kb of 3′ flanking sequences directs efficient and integration-site-independent expression in hematopoietic cells of myeloid lineages. Establishing such appropriate transcriptional regulation was significant in two ways. First, it provided evidence that this assay system, involving YAC DNA transfer, in vitro-differentiated ES cells, and expression analysis, is suitable for studying gene regulation in differentiating cells; second, these findings indicate that all the sequences necessary for high-level, tissue-restricted expression of this gene are contained within a single YAC clone.

Several studies have demonstrated efficient transgene expression through the transfer of YAC DNA into ES cells, suggesting that YACs can act as independent transcriptional units (21, 36, 54, 61, 65). Of particular relevance have been transgenic studies using YACs containing the human β-globin locus to examine the role of the LCR in controlling globin gene switching. Previous analyses of LCR function in regulating developmental control relied on small constructs with various portions of the β-globin gene cluster and limited amounts of gene-flanking sequences, often distorting the organization of the locus and spacing among distal regulatory elements (20, 46, 49). Significant differences in the expression profiles of LCR/globin transgenes were observed. YAC transgenes, however, provide an opportunity to study the locus in its "natural" context; indeed, transgenic mice harboring the β-globin YAC were found to express the different globin genes in a tissue-specific and developmentally correct manner, paralleling the pattern of expression observed from the endogenous mouse β-globin locus (21, 54).

Here, we have demonstrated that the introduction of a YAC clone recapitulates appropriate gp91-phox gene expression in the myeloid lineages of mouse ES cells. These findings stand in sharp contrast to those from prior transgenic experiments with elements derived from this locus. In those experiments, reporter constructs containing as much as 10 kb of flanking sequences failed to recapitulate normal patterns of expression, with respect to either the lineages in which transgenes were expressed or the levels at which they were expressed (62, 63). The ability of YAC Y1 to confer position-independent, high-level transgene expression suggests the presence of regulatory elements with both insulator and enhancer properties. To further investigate and identify these *cis* elements, we performed deletional analysis of the upstream sequences.

**Variegated transgene expression, a reflection of multiple regulatory elements required in concert.** The most striking finding in our deletion analysis is the marked variegation of expression at the single-cell level after the removal of upstream sequences. Recent studies have suggested that the phenomenon of variegation of expression is a consequence of alterations in chromatin structure, which in turn affects the probability of establishing an active transcriptional complex within a single cell (for a review, see reference 15). The commitment toward transcription appears to be a stochastic process, since at any given integration site a certain probability of transcription exists (69, 72). Transcriptional activity reflects the local chromatin environment at the site of integration of the transgene and the number and variety of *cis*-regulatory elements present within it. Given that gene expression is not often monitored at the single-cell level in transgenic experiments, the frequency of variegation in the expression of various gene constructs is unknown and may be greatly underestimated. Indeed, only recently has position-dependent variegation in transgenic mouse lines carrying a globin promoter-αHS-40 construct been reported (58).

Our deletion studies suggest that multiple regulatory elements that act in concert are required to establish and maintain an open chromatin domain sufficient for transcription of the human gp91-phox gene. The removal of flanking sequences far upstream, while not appreciably influencing the level of transcription in an "on" cell, adversely affects the number of cells capable of expressing the locus. These findings are compatible with models that suggest the need to recruit a critical, threshold set of regulatory factors to alter chromatin structure and overcome heterochromatic silencing (18). Through the analysis of deletions at the 5′ terminus of the gp91-phox locus, we observed that changes in the level of transgene expression were unpredictable and highly variable, an outcome attributable to random sites of integration and loss of important *cis*-regulatory elements. Transgenes with similar flanking sequences were expressed in a large number of cells. Overall, however, clones with larger 5′ flanking sequences expressed in a higher fraction of cells than did those with less flanking DNA.

Recent evidence has demonstrated the existence of DNA elements that protect transgene expression from chromosomal position effects (19, 40, 42, 47, 55, 59). Such regulators consist of a broad, loosely classified set of diverse and possibly distinct elements, including insulators, boundary elements, matrix attachment regions, and domains with chromatin opening activity. They generally do not activate reporter genes in traditional assays but function to define the ends of a chromatin domain. Insulators and boundary elements (as their names suggest) not only help to confine *cis*-regulatory elements to act within a restricted domain but also serve to insulate a domain from the repressive effects of heterochromatin (7). Domains with chromatin opening activity, on the other hand, act by propagating changes in chromatin structure and, as a consequence, increase the probability of forming a stable transcriptional complex. We suggest that the loss of such a class of elements contributes to a variegated phenotype but may not lead to complete loss of transgene expression. We further emphasize that these elements may not necessarily be marked by sites of strong DNase

I hypersensitivity and may lie significantly distal to regions possessing enhancer activity. Our deletion studies of clones with the most distal 5′ sequences removed provide evidence for the existence of important regulatory elements upstream of even the most distal hypersensitive site (HS IV). In addition, recent studies of the human CD2, adenosine deaminase, and β-globin LCRs have suggested the requirement for sequences which act in concert with core enhancer elements (1, 16, 35).

Progressively larger deletions from the 5′ terminus of Y1 lead to more drastic reductions in the level of transgene expression, presumably due to the loss of additional *cis*-regulatory elements. LCRs are known to be essential for directing high-level, tissue-specific expression. Enhancers are important components of LCRs and are thought to function with other *cis* elements to establish and maintain open chromatin domains (13, 16, 37, 70). Studies of enhancer function have suggested that enhancers act by increasing the probability of generating a stable transcriptional complex and that deletions of enhancer sequences reduce that chance of establishment (69, 72). Deletions of individual regulatory elements may not contribute significantly to loss of gene expression since functionally redundant elements may exist (17). However, when several elements are deleted, more pronounced changes in chromatin structure occur, resulting in major defects in gene expression. Our deletion studies, therefore, argue for cooperation between multiple regulatory elements upstream of the human gp91-phox locus. Only with the inclusion of all elements, such as those contained within YAC Y1, is high-level, tissue-restricted expression maintained in all cells.

Although the underlying mechanisms for transcriptional control are still not well understood, the remodeling of chromatin structure is believed to be an essential step. DNase I hypersensitivity formation has served as an indicator of active or potentially active regulatory elements (for reviews, see references 12 and 28). Several investigators have suggested that regulatory elements associated with hypersensitive sites are responsible for mediating factor accessibility, altering chromatin structure, and facilitating protein-protein interactions (13, 16, 27, 37, 38, 44). Within the 5′ flanking sequences of the gp91-phox locus, we found four strong, myeloid lineage-specific hypersensitive sites located at −13, −15, −28, and −29 kb. Although we have not directly tested their ability to enhance transcription, our data strongly implicate an important role for these regions in gp91-phox gene expression. We have shown that clones carrying HS I and HS II but lacking HS III and HS IV express poorly. However, despite this, expression is still detectable and remains restricted to myeloid cells. Thus, the proximal hypersensitive sites may confer tissue specificity, while the distal hypersensitive sites may provide enhancement. Alternatively, the proximal hypersensitive sites and the factors associated with them may be involved in maintaining, rather than establishing, an active transcriptional state.

Recent studies have suggested that hypersensitivity is generated and maintained through a stochastic mechanism, a phenomenon quite similar to the binary pattern of expression observed among our deletion clones. Detailed analysis of the chicken $\beta^A/\varepsilon$ globin gene enhancer found that erythroid factors binding at the enhancer contribute additively to open chromatin in an all-or-none fashion (5). Furthermore, mutations that reduced the number of sites for factor binding lowered the probability of factor accessibility at the enhancer. Surprisingly, the decrease in accessibility was manifested in a mixed population of nuclei which either did or did not display nuclease hypersensitivity. Studies of the chicken lysozyme locus also support a reduced efficiency for hypersensitive site formation when critical regulatory elements are deleted (32, 33). Position-independent transgene expression was found to correlate with proper hypersensitive site formation; in contrast, transgenes that displayed a variegated phenotype were greatly impaired in their ability to form hypersensitive sites. Taken together, these studies implicate a critical role for chromatin structure in regulating appropriate gene expression.

Our results highlight the utility of YACs for the study of gene regulation in a complex locus. We demonstrated that the genomic sequences contained within a 130-kb gp91-phox YAC are sufficient to drive position-independent, high-level, tissue-specific expression. Deletion analysis of YAC Y1 suggests the presence of multiple control elements that act in concert to provide full transcriptional activity. The removal of one or more of these elements impairs expression, most notably reflected in heterogeneity of expression at the single-cell level. Given that we have monitored protein expression from an intact coding region rather than from a bacterial reporter gene (such as *lacZ*), we suspect that variegated transgene expression in mammalian cells is much more common than hitherto appreciated and is not confined to the use of small plasmid constructs or specific reporter genes or the introduction of constructs into particular cell lineages (such as terminally differentiating erythrocytes). Furthermore, we propose that a combination of factors, including the effects of flanking heterochromatin and the lack of *cis*-acting control elements, contribute to transgene silencing. Our studies are consistent with the emerging view that chromatin structure plays a critical role in determining appropriate transcriptional regulation. The ability to assess the functions of regulatory elements within large DNA regions should provide significant insights into the molecular mechanisms that govern gene transcription in vivo.

## ACKNOWLEDGMENTS

We are grateful to Mike McDevitt and Karen Kotkow for helpful discussions, Tony Monaco for the gp91-phox YACs, Peter Newburger for initial characterizations of these YAC clones, Arthur Verhoeven and Dirk Roos for MAb48, and Philip Hieter for various yeast protocols, vectors, and host strains.

S.H.O. is an Investigator of the Howard Hughes Medical Institute. This work was supported by a grant from the National Institutes of Health.

## REFERENCES

1. **Aronow, B. J., C. A. Ebert, M. T. Valerius, S. S. Potter, D. A. Wiginton, D. P. Witte, and J. J. Hutton.** 1995. Dissecting a locus control region: facilitation of enhancer function by extended enhancer-flanking sequences. Mol. Cell. Biol. **15:**1123–1135.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1994. Current protocols in molecular biology. Greene Publishing Associates, New York, N.Y.
3. **Bonifer, C., M. Vidal, F. Grosveld, and A. E. Sippel.** 1990. Tissue-specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. EMBO J. **9:**2843–2848.
4. **Botstein, D., and G. R. Fink.** 1988. Yeast: an experimental organism for modern biology. Science **240:**1439–1443.
5. **Boyes, J., and G. Felsenfeld.** 1996. Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. EMBO J. **15:**2496–2507.
6. **Burgers, P. M. J., and K. J. Percival.** 1987. Transformation of yeast spheroplasts without cell fusion. Anal. Biochem. **163:**391–397.
7. **Chung, J. H., M. Whiteley, and G. Felsenfeld.** 1993. A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell **74:**505–514.
8. **Cockerill, P. N., and W. T. Garrard.** 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer on a region containing topoisomerase II sites. Cell **44:**273–282.
9. **Corces, V. G.** 1995. Keeping enhancers under control. Nature **376:**462–463.
10. **Cuoto, L. B., E. A. Spangler, and E. M. Rubin.** 1989. A method for the preparative isolation and concentration of intact yeast artificial chromosomes. Nucleic Acids Res. **20:**2693–2698.
11. **Dillon, N., and F. Grosveld.** 1994. Chromatin domains as potential units of eukaryotic gene function. Curr. Opin. Genet. Dev. **4:**260–264.

12. **Elgin, S. C. R.** 1988. The formation and function of DNase I hypersensitive sites in the process of gene activation. J. Biol. Chem. **263:**19259–19262.
13. **Ellis, J., K. C. Tan-un, A. Harper, D. Michalovich, N. Yannoutsos, S. Philipsen, and F. Grosveld.** 1996. A dominant chromatin-opening activity in 5′ hypersensitive site 3 of the human β-globin locus control region. EMBO J. **15:**562–568.
14. **Felsenfeld, G.** 1992. Chromatin as an essential part of the transcriptional mechanism. Nature **355:**219–224.
15. **Felsenfeld, G.** 1996. Chromatin unfolds. Cell **86:**13–19.
16. **Festenstein, R., M. Tolaini, P. Corbella, C. Mamalaki, J. Parrington, M. Fox, A. Miliou, M. Jones, and D. Kiousis.** 1996. Locus control region function and heterochromatin-induced position effect variegation. Science **271:** 1123–1125.
17. **Fiering, S., E. Epner, K. Robinson, Y. Zhuang, A. Telling, M. Hu, D. I. K. Martin, T. Enver, T. J. Ley, and M. Groudine.** 1995. Targeted deletion of 5′ HS2 of the murine β-globin LCR reveals that it is not essential for proper regulation of the b-globin locus. Genes Dev. **9:**2203–2213.
18. **Fiering, S., J. P. Northrop, G. P. Nolan, P. S. Mattila, G. R. Crabtree, and L. A. Herzenberg.** 1990. Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the T-cell antigen receptor. Genes Dev. **4:**1823–1834.
19. **Forrester, W. C., C. van Genderen, T. Jenuwein, and R. Grosschedl.** 1994. Dependence of enhancer-mediated transcription of the immunoglobulin μ gene on nuclear matrix attachment regions. Science **265:**1221–1225.
20. **Fraser, P., S. Pruzina, M. Antoniou, and F. Grosveld.** 1993. Each hypersensitive site of the human β-globin locus control region confers a different developmental pattern of expression of the globin genes. Genes Dev. **7:**106–113.
21. **Gaensler, K. M. L., M. Kitamura, and Y. W. Kan.** 1993. Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human β-globin locus in transgenic mice. Proc. Natl. Acad. Sci. USA **90:**11381–11385.
22. **Ganss, R., L. Montoliu, A. P. Monaghan, and G. Schutz.** 1994. A cell-specific enhancer far upstream of the mouse tyrosinase gene confers high-level and copy number-related expression in transgenic mice. EMBO J. **13:**3083–3093.
23. **Gasser, S. M., and U. K. Laemmli.** 1986. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D. melanogaster. Cell **46:**521–530.
24. **Gerasimova, T. I., and V. G. Corces.** 1996. Boundary and insulator elements in chromosomes. Curr. Opin. Genet. Dev. **6:**185–192.
25. **Geyer, P. K., and V. G. Corces.** 1992. DNA position-specific repression of transcription by a Drosophila Zn finger protein. Genes Dev. **6:**1865–1873.
26. **Greer, P., V. Maltby, J. Rossant, A. Bernstein, and P. Pawson.** 1990. Myeloid expression of the human c-*fps/fes* proto-oncogene in transgenic mice. Mol. Cell. Biol. **10:**2521–2527.
27. **Grewal, T., M. Theisen, U. Borgmeyer, T. Grussenmeyer, R. A. W. Rupp, A. Stief, F. Qian, A. Hecht, and A. E. Sippel.** 1992. The 6.1-kilobase chicken lysozyme enhancer is a multifactorial complex containing several cell-type-specific elements. Mol. Cell. Biol. **12:**2339–2350.
28. **Gross, D. S., and W. T. Gerrard.** 1988. Nuclease hypersensitive sites in chromatin. Annu. Rev. Biochem. **57:**159–197.
29. **Grosschedl, R., and M. Marx.** 1988. Stable propagation of the active transcriptional state of an immunoglobulin gene requires continuous enhancer function. Cell **55:**645–654.
30. **Grosveld, F., G. B. van Assendelft, D. R. Greaves, and B. Kollias.** 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell **51:**975–985.
31. **Hermanson, G. G., M. F. Hoekstra, D. L. McElligott, and G. A. Evans.** 1991. Rescue of end fragments of yeast artificial chromosomes by homologous recombination in yeast. Nucleic Acid Res. **19:**4943–4948.
32. **Huber, M. C., F. X. Bosch, A. E. Sippel, and C. Bonifer.** 1994. Chromosomal position effects in chicken lysozyme gene transgenic mice are correlated with suppression of DNase I hypersensitive site formation. Nucleic Acids Res. **22:**4195–4201.
33. **Huber, M. C., G. Kruger, and C. Bonifer.** 1996. Genomic position effects lead to an inefficient reorganization of nucleosomes in the 5′-regulatory region of the chicken lysozyme locus in transgenic mice. Nucleic Acids Res. **24:**1443–1452.
34. **Huxley, C., Y. Hagino, D. Schlessinger, and M. V. Olson.** 1991. The human HPRT gene on a yeast artificial chromosome is functional when transferred to mouse cells by cell fusion. Genomics **9:**742–750.
35. **Jackson, J. D., H. Petrykowska, S. Philipsen, W. Miller, and R. Hardison.** 1996. Role of DNA sequences outside of the cores of DNase hypersensitive sites (HSs) in function of β-globin locus control region. J. Biol. Chem. **271:**11871–11878.
36. **Jakobovits, A., A. L. Moore, L. L. Green, G. J. Vergara, C. E. Maynard-Currie, H. A. Austin, and S. Klapholz.** 1993. Germ-line transmission and expression of a human-derived yeast artificial chromosome. Nature **362:**255–258.
37. **Jenuwein, T., W. Forrester, R.-G. Qui, and R. Grosschedl.** 1993. The immunoglobulin μ enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. Genes Dev. **7:**2016–2032.
38. **Jimenez, G., K. B. Gale, and T. Enver.** 1992. The mouse β-globin locus control region: hypersensitive sites 3 and 4. Nucleic Acids Res. **20:**5797–5803.
39. **Jones, B. K., B. R. Monks, S. A. Liebhaber, and N. E. Cooke.** 1995. The human growth hormone gene is regulated by a multicomponent locus control region. Mol. Cell. Biol. **15:**7010–7021.
40. **Kalos, M., and R. E. K. Fournier.** 1995. Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. Mol. Cell. Biol. **15:**198–207.
41. **Keller, G., M. Kennedy, T. Papayannopoulou, and M. V. Wiles.** 1993. Hematopoietic commitment during embryonic stem cell differentiation in culture. Mol. Cell. Biol. **13:**472–486.
42. **Kellum, R., and P. Schedl.** 1992. A group of scs elements function as domain boundaries in an enhancer-blocking assay. Mol. Cell. Biol. **12:**2424–2431.
43. **Kellum, R., and P. Schedl.** 1991. A position-effect assay for boundaries of higher order chromosomal domains. Cell **64:**941–950.
44. **Lake, R. A., D. Wotton, and M. J. Owen.** 1990. A 3′ transcriptional enhancer regulates tissue-specific expression of the human CD2 gene. EMBO J. **9:**3129–3136.
45. **Lewin, B.** 1994. Chromatin and gene expression: constant questions, but changing answers. Cell **79:**397–406.
46. **Lloyd, J. A., J. M. Krakowsky, S. C. Crable, and J. B. Lingrel.** 1992. Human γ- to β-globin gene switching using a mini construct in transgenic mice. Mol. Cell. Biol. **12:**1561–1567.
47. **McKnight, R. A., A. Shamay, L. Sankaran, R. J. Wall, and L. Hennighausen.** 1992. Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice. Proc. Natl. Acad. Sci. USA **89:**6943–6947.
48. **Mizushima, S., and S. Nagata.** 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. **18:**5322.
49. **Morley, B. J., C. A. Abbott, J. A. Sharpe, J. Lida, P. S. Chan-Thomas, and W. G. Wood.** 1992. A single β-globin locus control region element (5′ hypersensitive site 2) is sufficient for developmental regulation of human globin genes in transgenic mice. Mol. Cell. Biol. **12:**2057–2066.
50. **Mortensen, R. M., D. A. Conner, S. Chao, A. A. T. Geisterfer-Lowrance, and J. G. Seidman.** 1992. Production of homozygous mutant ES cells with a single targeting construct. Mol. Cell. Biol. **12:**2391–2395.
51. **Moyzis, R. K., D. C. Torney, J. Meyne, J. M. Buckingham, J.-R. Wu, C. Burks, K. M. Sirotkin, and W. B. Goad.** 1989. The distribution of interspersed repetitive DNA sequences in the human genome. Genomics **4:**273–289.
52. **Pavan, W. J., P. Hieter, and R. H. Reeves.** 1990. Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. Proc. Natl. Acad. Sci. USA **87:**1300–1304.
53. **Pavan, W. J., P. Hieter, D. Sears, A. Burkhoff, and R. H. Reeves.** 1991. High-efficiency yeast artificial chromosome fragmentation vectors. Gene **106:**125–127.
54. **Peterson, K. R., C. H. Clegg, C. Huxley, B. M. Josephson, H. S. Haugen, T. Furukawa, and G. Stamatoyannopoulos.** 1993. Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human β-globin locus display proper developmental control of human globin genes. Proc. Natl. Acad. Sci. USA **90:**7593–7597.
55. **Phi-van, L., J. P. von Kries, W. Ostertag, and W. H. Stratling.** 1990. The chicken lysozyme 5′ matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. Mol. Cell. Biol. **10:**2302–2307.
56. **Reitman, M., E. Lee, H. Westphal, and G. Felsenfeld.** 1993. An enhancer/locus control region is not sufficient to open chromatin. Mol. Cell. Biol. **13:**3990–3998.
57. **Robertson, E.** 1987. Teratocarcinomas and embryonic stem cells: a practical approach. IRL Press, Oxford, United Kingdom.
58. **Robertson, G., D. Garrick, W. Wu, M. Kearns, D. Martin, and E. Whitelaw.** 1995. Position-dependent variegation of globin transgene expression in mice. Proc. Natl. Acad. Sci. USA **92:**5371–5375.
59. **Roseman, R. R., V. Pirotta, and P. K. Geyer.** 1993. The su(Hw) protein insulates expression of the Drosophila melanogaster white gene from chromosomal position-effects. EMBO J. **12:**435–442.
60. **Royer-Pokora, B., L. M. Kunkel, A. P. Monaco, S. C. Goff, P. E. Newburger, R. L. Baehner, F. S. Cole, J. T. Curnutte, and S. H. Orkin.** 1986. Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. Nature **322:**32–38.
61. **Schedl, A., L. Montoliu, G. Kelsey, and G. Schutz.** 1993. A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice. Nature **362:**258–261.
62. **Skalnik, D., D. M. Dorfman, A. S. Perkins, N. A. Jenkins, N. G. Copeland, and S. H. Orkin.** 1991. Targeting of transgene expression to monocyte/macrophages by the gp91-phox promoter and consequent histiocytic malignancies. Proc. Natl. Acad. Sci. USA **88:**8505–8509.
63. **Skalnik, D., D. M. Dorfman, D. Williams, and S. H. Orkin.** 1991. Restriction of neuroblastoma to the prostate gland in transgenic mice. Mol. Cell. Biol. **11:**4518–4527.

64. **Stief, A., D. M. Winter, W. H. Stratling, and A. E. Sippel.** 1989. A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature **341:**343–345.
65. **Strauss, W. M., J. Dausman, C. Beard, J. B. Lawrence, and R. Jaenisch.** 1993. Germ line transmission of a yeast artificial chromosome spanning the murine α1(I) collagen locus. Science **259:**1904–1907.
66. **Tybulewicz, V. L. J., C. E. Crawford, P. K. Jackson, R. T. Bronson, and R. C. Mulligan.** 1991. Neonatal lethality and lymphopenia in mice with a homozygous disruption of the c-abl proto-oncogene. Cell **65:**1153–1163.
67. **Verhoeven, A. J., B. G. J. M. Bolscher, L. J. Meerhof, R. van Zwieten, J. Keijer, R. S. Weening, and D. Roos.** 1989. Characterization of two monoclonal antibodies against cytochrome b558 of human neutrophils. Blood **73:**1686–1694.
68. **Vollrath, D., R. W. Davis, C. Connelly, and P. Hieter.** 1988. Physical mapping of large DNA by chromosome fragmentation. Proc. Natl. Acad. Sci. USA **85:**6027–6031.
69. **Walters, M. C., S. Fiering, J. Eidemiller, W. Magis, M. Groudine, and D. I. K. Martin.** 1995. Enhancers increase the probability but not the level of gene expression. Proc. Natl. Acad. Sci. USA **92:**7125–7129.
70. **Walters, M. C., W. Magis, S. Fiering, J. Eidemiller, D. Scalzo, M. Groudine, and D. I. K. Martin.** 1996. Transcriptional enhancers act in cis to suppress position-effect variegation. Genes Dev. **10:**185–195.
71. **Wang, Y., J. P. Macke, S. L. Merbs, D. J. Zack, B. Klaunberg, J. Bennett, J. Gearhart, and J. Nathans.** 1992. A locus control region adjacent to the human red and green visual pigment genes. Neuron **9:**429–440.
72. **Weintraub, H.** 1988. Formation of stable transcription complexes as assayed by analysis of individual templates. Proc. Natl. Acad. Sci. USA **85:**5819–5823.
73. **Wiles, M. V., and G. Keller.** 1991. Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. Development **111:**259–267.