THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 270, No. 30, Issue of July 30, pp. 18019–18025, 1995
Printed in U.S.A.

Structure, Expression, and Chromosomal Assignment of the Human Gene Encoding Nuclear Respiratory Factor 1*

(Received for publication, March 7, 1995, and in revised form, May 26, 1995)

Lekha Gopalakrishnan and Richard C. Scarpulla‡

From the Department of Cell and Molecular Biology, Northwestern University Medical School, Chicago, Illinois 60611

Nuclear respiratory factor 1 (NRF-1) is a transcription factor that acts on nuclear genes encoding respiratory subunits and components of the mitochondrial transcription and replication machinery. Here we describe the isolation and characterization of the human gene encoding NRF-1. The human genomic sequences detected with NRF-1 cDNA probes at high stringency are all contained within seven overlapping recombinant λ clones. The NRF-1 gene encompassed by these recombinants spans ~65 kilobases (kb) and has 11 exons and 10 introns that range in size from 0.8 to 15 kb. A rapid amplification of cDNA ends-polymerase chain reaction product containing the 5′-terminus of the NRF-1 cDNA has two exons from the 5′-untranslated region and terminates at a major transcription initiation site identified by S1 nuclease mapping. A genomic fragment containing a portion of the 5′-terminal exon and an additional 1 kb upstream had a functional promoter that was active in transfected COS cells, HeLa cells, and L6 myoblasts. The transcription initiation site utilized by the transfected promoter corresponded to that used by the endogenous gene *in vivo*. NRF-1 mRNA was expressed at very low levels in rat tissues compared with cytochrome *c* and, unlike cytochrome *c*, was most abundantly expressed in lung and testis. The NRF-1 gene was localized to human chromosome 7 by analysis of DNA from a panel of human-hamster cell hybrids with human-specific NRF-1 polymerase chain reaction primers. This assignment was further refined to 7q31 by cohybridization of NRF-1- and chromosome 7-specific probes to human metaphase chromosomes. These analyses should be useful in evaluating the potential role of NRF-1 in mitochondrial diseases resulting from defects in the nuclear control of mitochondrial function.

The electron transport chain and oxidative phosphorylation system rely upon the functional interplay of gene products expressed from both nuclear and mitochondrial genetic systems. Because of the limited coding capacity of the mitochondrial genome, nuclear genes must provide the majority of the respiratory subunits and all of the gene products necessary for mtDNA[1] transcription and replication (1, 2). Several of the latter have been characterized from vertebrates in recent years. These include a mitochondrial RNA-processing endonuclease that has been implicated in mtDNA heavy strand replication (1), a mitochondrial termination factor that may contribute to the proper rRNA and mRNA stoichiometries within the mitochondria (3), and a mitochondrial transcription factor (mtTFA) that is required for efficient heavy and light strand transcription (1, 4). The latter may also participate in the replication and maintenance of mtDNA (1, 4, 5). Regulatory mechanisms that coordinate the expression of these gene products with the expression of respiratory subunits may, in part, serve to integrate nuclear and mitochondrial genetic systems.

A transcriptional analysis of nuclear cytochrome *c* (6, 7) and cytochrome oxidase subunit (8, 9) genes was undertaken to identify regulatory factors that might serve such an integrative function. These studies led to the identification, purification, and molecular cloning of nuclear respiratory factors NRF-1 (10, 11) and NRF-2 (9, 12). Functional recognition sites for one or both of these nuclear transcription factors reside in the promoters of nuclear genes that encode many of the respiratory subunits, mtDNA transcription and replication factors, and the rate-limiting heme biosynthetic enzyme (10, 11, 13, 14). These findings led to the hypothesis that NRFs may facilitate nuclear-mitochondrial interactions through their activation of these nuclear target genes (11, 13).

Recently, a fatal human genetic defect resulting in the depletion of mtDNA has been described (15, 16). Affected individuals have severely reduced levels of mtDNA in affected tissues, leading to a diminished respiratory phenotype. The majority of known mitochondrial diseases are associated with mutations in the mitochondrial genome and can be genetically transmitted through maternal inheritance (2). By contrast, mtDNA depletion appears to be inherited as a nuclear gene defect (15), thus implicating a component of the mtDNA replication machinery itself or a regulatory gene required for the maintenance of mtDNA. It is intriguing in this context that mtTFA levels have been shown to vary with mtDNA in patients displaying the extremes of depletion and pathological proliferation of mtDNA (17, 18) and that NRF-1 has been found to be an important activator of the mtTFA promoter (13).

Here, we describe the structure, expression, and chromosomal assignment of the human gene encoding NRF-1. This information should be useful in elucidating the potential involvement of the NRF-1 gene in human mitochondrial diseases resulting from nuclear gene defects.

EXPERIMENTAL PROCEDURES

Isolation and Sequencing of Genomic NRF-1 Clones—A human placenta genomic library constructed from *Sau*3AI partial digestion products cloned into EMBL-3 SP6/T7 (CLONTECH) was screened with

* This work was supported in part by United States Public Health Service Grant GM32525-12 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by Faculty Research Award FRA-361 from the American Cancer Society. To whom correspondence should be addressed: Dept. of Cell and Molecular Biology, Northwestern University Medical School, 303 East Chicago Ave., Chicago, IL 60611. Tel.: 312-503-2946; Fax: 312-503-0798.

[1] The abbreviations used are: mtDNA, mitochondrial DNA; mtTFA, mitochondrial transcription factor A; NRF, nuclear respiratory factor; RACE, rapid amplification of cDNA ends; PCR, polymerase chain reaction; CAT, chloramphenicol acetyltransferase; kb, kilobase(s); bp, base pair(s); UTR, untranslated region.

random primer-labeled radioactive probes (19) derived from the NRF-1 cDNA (11) as described previously (20). Phage DNA was prepared from purified positive isolates by an established procedure (21). Hybridizing fragments were first cloned into pGEM7zf(+) or pGEM5zf(+) (Promega) and then subcloned into M13 vectors (New England Biolabs Inc.) for DNA sequencing on both strands by the dideoxy chain termination method (22).

Genomic Blotting—Human genomic DNA (10 μg) was digested with *Pst*I; electrophoresed on a 1% agarose gel along with *Pst*I-digested DNA from ϕ 9-1, ϕ 47-1, ϕ 39-1, ϕ 6-2, and ϕ 12-1; and transferred to nitrocellulose. The filter was probed with a random primer-labeled (19) NRF-1 cDNA probe containing the entire coding region (11) as described previously (23). Hybridization was carried out as described (23), except that the filter was washed consecutively in 2 × SSC, 0.5% SDS and in 2 × SSC, 0.1% SDS at room temperature and in 0.1 × SSC, 0.5% SDS for 30 min at 65 °C.

RACE-PCR—The following primers were used for the amplification of the 5′-end of NRF-1 by the method of RACE-PCR: primer 1 (cDNA synthesis), 5′-CCACTGCATGTGCTTCTATGGTAG-3′; and primer 2 (nested primer for PCR amplification), 5′-CACTCCGTGTTCCTC-CATG-3′. RACE-PCR was carried out with reagents provided in the 5′-Amplifinder RACE kit obtained from CLONTECH according to the manufacturer's instructions. Briefly, cDNA was synthesized from 2 μg of HeLa poly(A)$^+$ RNA using a primer within the NRF-1 coding region extending from positions 263 to 286 in an antisense direction (primer 1). Amplification was performed with a "nested" primer (primer 2) complementary to the NRF-1 coding region (positions 225–243 in an antisense direction) and an anchor primer provided with the kit. PCR mixtures were heated at 82 °C for 1 min prior to addition of primers. After addition of primers, amplification was carried at 94 °C for 45 s, 60 °C for 45 s, and 72 °C for 2 min for 35 cycles with a final extension time of 7 min at 72 °C. The PCR product was extracted once with chloroform/isoamyl alcohol (24:1) and precipitated with ethanol. A portion of the PCR product was cloned into pGEM-T (Promega), and the insert was sequenced on both strands by the dideoxy chain termination method (22).

S1 Nuclease Mapping—Uniformly labeled, single-stranded probes were generated by the method of Burke (24) using an M13 template containing sequences upstream of position 91 and the antisense primer shown in Fig. 3*A*. Hybridization to RNA and S1 nuclease digestion were carried out as described previously (25).

Promoter Plasmids and CAT Assays—A 2-kb *Hin*dIII fragment containing 5′UTR-1 and upstream sequences was excised from ϕ 2-1 and cut at the *Nae*I site at position +38. The resulting 1.1-kb fragment was cloned into the *Hin*dIII and *Sma*I sites in the polylinker of pGEM7zf(+). An *Nsi*I-*Asp*718I fragment was excised from this construction and cloned into the corresponding sites in the polylinker of RC4CAT/+17 (9), thus fusing the first 38 bp of 5′UTR-1 to position +17 of the RC4CAT first exon to generate NRF-1CAT.

COS-1, HeLa, and L6 cells were cultured in Dulbecco's modified Eagle's medium containing 10% (v/v) bovine calf serum, 100 units/ml penicillin, and 100 μg/ml streptomycin. COS-1 and HeLa cells were allowed to grow to 80% confluence prior to transfections, while L6 cells were grown to ~50% confluence in 100-mm dishes. Transfections were performed in 100-mm plates by the $CaPO_4$ precipitation method as previously described (8) using 20 μg of DNA consisting of 5 μg of test plasmid, 1 μg of pRSVLacZ (26), and 14 μg of pGEM-4 blue carrier plasmid. The $CaPO_4$ precipitate was prepared by the dropwise addition of 60 μg of DNA (in the proportions described above) and 180 μl of 2 M $CaCl_2$ to a 15-ml polypropylene tube containing 1.5 ml of 208 mM HEPES, 136 mM NaCl, 5 mM KCl, 11.2 mM glucose, and 1.4 mM Na_2HPO_4 (pH 7.1) and water to give a final volume of 3 ml. Addition of DNA was accompanied by the bubbling of air into the mixture in order to facilitate the formation of a uniform precipitate. The precipitate was incubated at room temperature for 30 min with vortexing every 10 min. $CaPO_4$/DNA precipitate (1 ml) was added to each of three 100-mm plates of cells. Twenty-four hours following transfection, cells were washed twice with 5 ml of phosphate-buffered saline and then refed. After an additional 24 h, cells from triplicate plates were harvested into 5 ml of phosphate-buffered saline. CAT assays were performed as described previously (27), and cell extracts were assayed for β-galactosidase activity (28) to normalize for transfection efficiency.

RNA Preparation and Ribonuclease Protection Assays—Total tissue RNA was prepared from adult Sprague-Dawley rats (Harlan Bioproducts for Science, Inc.) by LiCl precipitation as described (29). Equal amounts of RNA were analyzed for NRF-1 and cytochrome *c* message levels by RNase protection assays (30). The cytochrome *c* transcript was assayed with a SP6-generated riboprobe derived from a cloned 167-bp fragment spanning the splice junction between the rat cytochrome *c* first intron and second exon. The NRF-1 antisense riboprobe was generated from a cloned 305-bp fragment containing CE-8 of the rat NRF-1 gene and some flanking intron sequences. Hybridization of probe and RNA was carried out as described (6), except for a reduction in the RNase digestion time from 1 h to 30 min.

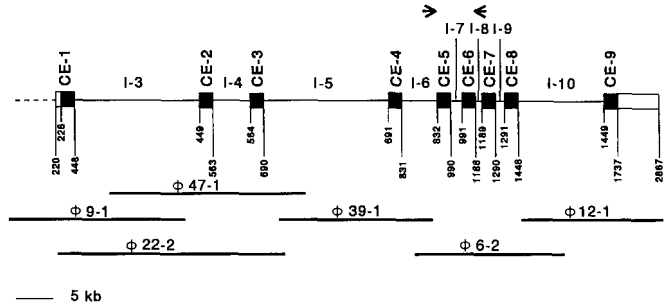

FIG. 1. **Genomic organization of the human NRF-1 gene.** The diagram depicts the intron/exon structure of the human NRF-1 gene with coding regions (*closed boxes*), 5′- and 3′-untranslated regions (*open boxes*), and introns (*solid line*). The coding exons (CE-1 through CE-9) and introns (I-3 through I-10) are numbered consecutively. The coordinates at the 5′- and 3′-ends of each exon indicate their position within the cDNA (mRNA) relative to the NRF-1 transcription initiation site (see Figs. 3 and 4). Approximate intron sizes are as follows: I-3, 15 kb; I-4, 5 kb; I-5, 15 kb; I-6, 5 kb; I-7, 1.2 kb; I-8, 0.8 kb; I-9, 1.3 kb; and I-10, 11 kb. The phage isolates (ϕ) containing genomic sequences and the extent of their overlap are indicated at the bottom (*bold solid lines*). The *arrows* flanking CE-5 and CE-6 indicate the approximate positions of the PCR primers discussed for Fig. 7. The *dashed line* upstream of CE-1 represents 5′-untranslated region and intron sequences contained within ϕ 9-1 as described for Fig. 3*B*. The sequences of CE-1 through CE-9 and their respective 5′- and 3′-splice junctions have been deposited in GenBank™ under accession numbers U18375–U18383. Exons are not drawn to scale.

PCR Amplification of Somatic Cell Hybrids—The following oligonucleotide primers were used: NRF-1 6-2F, 5′-GATGACCACACTGT-TCTCTTCC-3′; and NRF-1 6-2R, 5′-CTGGTAGCCCTCAGGTTTACTC-3′. NRF-1 6-2F is a sense primer specific for intron sequences immediately preceding the 5′-splice junction of CE-5, while NRF-1 6-2R is an antisense primer specific for the region downstream of the 3′-splice junction of CE-6 (see Fig. 1).

Purified genomic DNA from Chinese hamster ovary cells and a panel of human-hamster hybrids (see Table II) from BIOS Laboratories (New Haven, CT) were subjected to PCR amplification using primers NRF-1 6-2F and NRF-1 6-2R. Human (Sigma) and Chinese hamster ovary genomic DNAs (100 ng each) along with 250 ng of DNA from each of the hybrid lines were amplified in 100-μl reactions containing 40 pmol of each primer, 2.5 units of *Taq* polymerase, 200 μM each dNTP, 10 mM Tris, 50 mM KCl, 1.5 mM $MgCl_2$, and 0.001% (w/v) gelatin. After an initial denaturation at 95 °C for 5 min, *Taq* polymerase was added at 72 °C. Amplification was carried out at 95 °C for 1 min, 65 °C for 1 min, and 72 °C for 1 min for 35 cycles, followed by a 15-min extension at 72 °C. A 20-μl aliquot of each PCR was electrophoresed on a 1% agarose gel, and the products were visualized by ethidium bromide staining.

Fluorescence in Situ Hybridization—Chromosomal localization of the NRF-1 locus by fluorescence *in situ* hybridization was performed at BIOS Laboratories. Cesium chloride-banded DNA from ϕ 47-1 (see Fig. 1) was nick-translated with biotin-dUTP, combined with sheared human DNA, and hybridized to normal metaphase chromosomes derived from phytohemagglutinin-stimulated peripheral blood lymphocytes. Cohybridizations with a centromere-specific probe were performed to confirm the identity of the labeled chromosome. Hybridization signals were detected by incubation of the post-hybridization slides in fluorescein-conjugated avidin. Upon detection of the signal, the slides were counterstained with propium iodide and analyzed.

RESULTS

Isolation and Structural Organization of the Human NRF-1 Gene—Screening of a human placenta genomic library with a probe derived from the human NRF-1 cDNA clone (11) led to the initial isolation of six overlapping phages spanning ~57 kb of genomic DNA (Fig. 1). Regions of overlap were determined by restriction enzyme mapping and hybridization analysis of

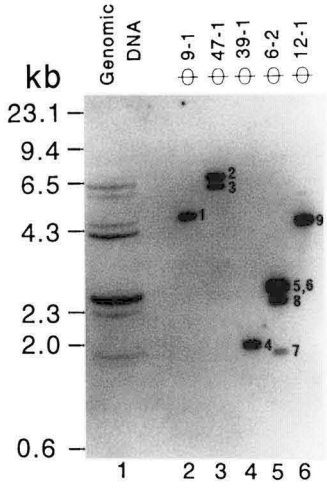

FIG. 2. **Hybridization analysis of NRF-1 restriction fragments in genomic DNA and in recombinant λ clones.** Human genomic DNA (*lane 1*) and DNA isolated from the indicated recombinant λ clones (*lanes 2–6*) were digested with *Pst*I and subjected to hybridization analysis as described under "Experimental Procedures." *Numbers 1–9* to the right of each hybridizing fragment refer to CE-1 through CE-9 depicted in Fig. 1. Size markers at left are from *Hin*dIII-digested λ DNA.

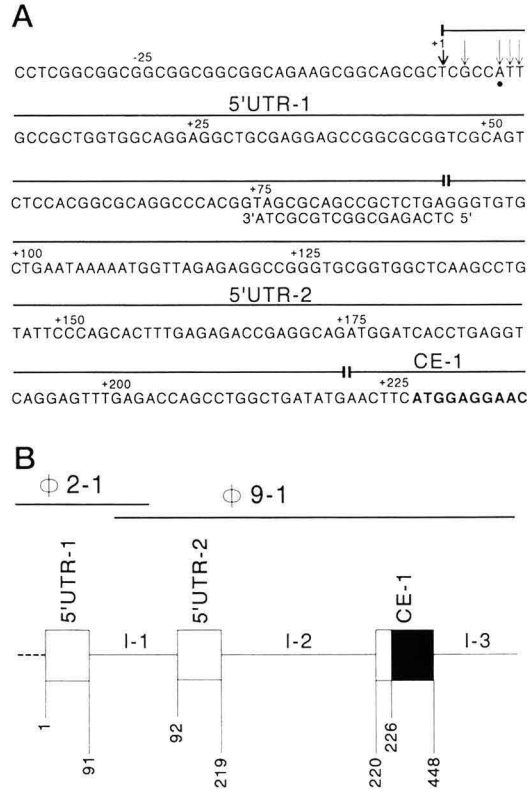

FIG. 3. **Analysis of the 5′-terminus of the NRF-1 cDNA by RACE-PCR and genomic organization of the 5′-terminal exons of the NRF-1 gene.** *A*, nucleotide sequence of the 5′-untranslated region of the NRF-1 cDNA cloned by RACE-PCR (GenBank™ accession number U27701) as described under "Experimental Procedures." *Solid lines* above the sequence mark the positions of two 5′-untranslated region exons (5′UTR-1 and 5′UTR-2) and the first coding exon (CE-1). *Numbers* refer to nucleotide positions relative to the transcription initiation sites (*arrows*) as determined by S1 nuclease analysis (see Fig. 4). The *filled circle* at position +6 indicates the first nucleotide of the RACE-PCR product. Translated frame is in *boldface letters*. *B*, diagram of the intron/exon organization of the 5′-terminus of the NRF-1 gene including the coding region (*closed box*), 5′-untranslated sequences (*open boxes*), introns (*solid lines*), and the 5′-flanking region (*dashed line*). The coordinates at the 5′- and 3′-ends of each exon indicate their position within the cDNA (mRNA) relative to the NRF-1 transcription initiation site (*A*; see Fig. 4). Regions encompassed by the overlapping genomic λ clones φ 2-1 and φ 9-1 are indicated at the top (*solid boldface lines*). Approximate intron sizes are as follows: I-1, 2.5 kb; and I-2, 5 kb.

phage DNA (data not shown). The DNA sequences were determined for subcloned segments from each isolate that hybridized to the NRF-1 cDNA and included coding regions, untranslated regions, and intron/exon junctions. These sequences have been deposited in GenBank™ under accession numbers U18375–U18383. The human gene consists of nine protein coding exons (CE-1 through CE-9) interrupted by introns ranging in size from 0.8 to 15 kb (Fig. 1).

Genomic hybridization analysis performed under normal stringency conditions using a NRF-1 cDNA probe revealed eight prominent *Pst*I restriction enzyme fragments in human genomic DNA (Fig. 2, *lane 1*). To determine whether the six phage isolates can account for these genomic fragments, DNA from each phage was digested with *Pst*I and run simultaneously. The results demonstrate a one-to-one correspondence between the NRF-1 *Pst*I fragments in genomic DNA and those present in the phage isolates (compare *lane 1* with *lanes 2–6*). Moreover, the relative intensities of the hybridization signals were very similar between genomic and phage DNAs. Although sequences more distantly related to NRF-1 may be present in the human genome, these results demonstrate that those most closely related to the NRF-1 cDNA are likely present in a single contiguous locus.

At its 5′-end, CE-1 has only 6 bp of 5′-untranslated region matching that of the original NRF-1 cDNA. The point of divergence with this cDNA coincides with the consensus for an intron acceptor splice junction, indicating that additional upstream exon(s) contain the remainder of the 5′-untranslated region. RACE-PCR was performed on HeLa cell poly(A)$^+$ RNA for the purpose of obtaining a complete 5′-untranslated region that would facilitate the identification and isolation of the 5′-end of the NRF-1 gene. Hybridization of the RACE-PCR product (Fig. 3*A*) to φ 9-1 revealed a single fragment containing a 5′-untranslated region exon designated 5′UTR-2 (Fig. 3, *A* and *B*). This exon was separated from CE-1 by an intron of ~5 kb (Fig. 3*B*, *I-2*). The remainder of the RACE-PCR product was not present on φ 9-1. However, rescreening the library with a synthetic oligomer containing a sequence within the RACE-PCR product just upstream from 5′UTR-2 yielded φ 2-1, which overlapped the 5′-end of φ 9-1 (Fig. 3*B*). An additional exon, designated 5′UTR-1, was present on a single restriction fragment from φ 2-1. This exon was colinear with the remainder of the RACE-PCR product, and no acceptor splice junction consensus was present in the genomic sequence coinciding with the 5′-RACE-PCR end. Therefore, 5′UTR-1 and 5′UTR-2 are contiguous with the NRF-1 gene and likely represent the 5′-terminal exons.

Identification of the NRF-1 Gene Promoter and Transcription Initiation Sites—S1 nuclease mapping was performed using total HeLa cell RNA to determine whether the 5′-ends of 5′UTR-1 and the RACE-PCR product coincide with a transcription initiation site. As shown in Fig. 4, S1 nuclease cleavage revealed a heterogeneous cluster of initiation sites within 8 bp. These initiation sites overlapped the 5′-end of the RACE-PCR product, and one coincided precisely with the 5′-terminal nucleotide (Fig. 3*A*).

If the sites revealed by S1 nuclease mapping are authentic initiation sites, the 5′-flanking DNA adjacent to these sites should contain promoter activity. To test this possibility, a restriction fragment containing ~1.1 kb of the 5′-untranslated region and 38 bp of 5′UTR-1 was cloned into a rat cytochrome *c*/CAT expression vector. In this construct, 5′UTR-1 is fused at

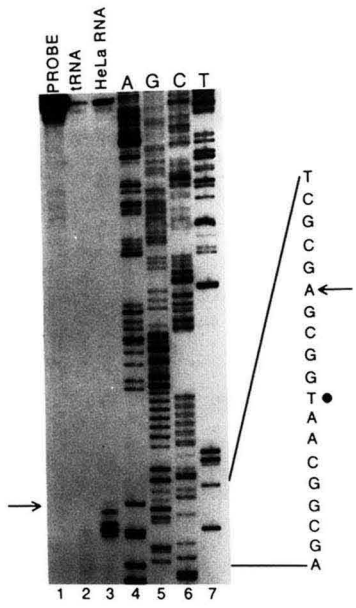

FIG. 4. **Mapping of the NRF-1 transcription initiation site.** A uniformly labeled, single-stranded antisense probe was generated from an M13 template containing a 1.1-kb NRF-1 genomic fragment extending upstream from the priming site at the 3′-terminus of 5′UTR-1 (see Fig. 3*A*). Yeast tRNA (*lanes 1* and *2*) and total HeLa RNA (*lane 3*) (20 μg each) were hybridized to the probe, and the hybrids were digested with 0 (*lane 1*) or 500 (*lanes 2* and *3*) units of S1 nuclease for 60 min. The products were analyzed on a 6% urea-acrylamide gel adjacent to a sequencing ladder (*lanes 4–7*) generated by extension of the same primer. Note that the sequence indicated alongside the sequencing ladder is the complement of the gene sequence described for Fig. 3*A*. The transcription initiation sites (*arrows*) and the 5′-terminal nucleotide of the RACE-PCR product (*filled circle*) are indicated.

nucleotide +38 to position +17 of the cytochrome *c* gene 5′-untranslated region (9). Thus, any NRF-1 promoter activity within this fragment would be expected to initiate transcription through the NRF-1 start sites revealed by S1 nuclease mapping. As shown in Table I, significant levels of CAT activity above the promoterless control resulted from transfection of the fusion gene into COS cells, HeLa cells, and L6 myoblasts. The NRF-1 promoter activity is 18–25-fold higher than the negative control RC4CAT/+17, but 6–11-fold lower than the cytochrome *c* promoter activity under identical conditions of transfection. The relative differences in these transfected promoters are consistent with the low steady-state level of NRF-1 mRNA compared with that of cytochrome *c* mRNA (see Fig. 6).

To verify that the transfected promoter utilizes the same initiation sites as the endogenous human gene, promoter-dependent 5′-end formation was assayed by RNase protection in transfected COS cells and L6 myoblasts. The riboprobe used in this assay is specific for the detection of primate NRF-1 genes (Fig. 5). No transcript of the size expressed from the endogenous NRF-1 gene in HeLa cells (Fig. 5, *lanes 1* and *6*) was detected in mock-transfected L6 myoblasts (*lane 7*) under conditions where it is readily detected in mock-transfected COS cells (*lane 2*). Transfection of the fusion gene into the rat L6 cells yielded a transcript of identical size to that expressed in the primate HeLa or COS cells (*lane 8*). Likewise, the level of NRF-1 transcript was elevated by transfection of the fusion gene into COS cells (*lane 3*). These results establish that 1) both the endogenous and transfected NRF-1 genes initiate transcription at the identical positions; 2) the cloned fragment contains at least a portion of the NRF-1 promoter utilized *in vivo*; and 3) 5′UTR-1 is the 5′-terminal exon of the human NRF-1 gene.

Tissue Distribution of NRF-1 mRNA Expression and Assignment of the NRF-1 Gene to Human Chromosome 7q31—In light of the proposed role for NRF-1 in nuclear-mitochondrial interactions, it was of interest to determine the tissue distribution of NRF-1 mRNA expression. As expected for a regulatory factor, the NRF-1 transcript is expressed at very low abundance in all cells and tissues tested. Therefore, to assay its relative distribution in rat tissues, a portion of the rat NRF-1 gene that is homologous to human CE-8 was cloned and sequenced for use in a sensitive riboprobe assay. The sequence of this fragment has been deposited in GenBank™ under accession number U27700. Cytochrome *c* transcripts, whose levels of expression are well established in rat tissues (31), were measured simultaneously as an internal control. It is evident from Fig. 6*A* that NRF-1 transcripts are barely detectable under conditions where cytochrome *c* transcripts are easily visualized. The

TABLE I
Comparison of NRF-1 and cytochrome c promoter activities in transfected cells

Reporter construction[a]	Relative CAT activity[b]		
	COS	HeLa	L6
RC4CAT/+17	1.0	1.0	1.0
NRF-1CAT	24.7 ± 4.8	20.3 ± 9.5	16.9 ± 3.5
RC4CAT/−326	282.6 ± 42.4	122.2 ± 3.9	110 ± 7.7

[a] Reporter plasmids were constructed from a derivative of the rat cytochrome *c* gene (RC4) in which the cytochrome *c* coding region was replaced by CAT (RC4CAT) as described under "Experimental Procedures." Coordinates indicated are relative to the cytochrome *c* transcription initiation sites.

[b] Reporter constructs (5 μg) were transfected into the indicated cell line along with pRSVLacZ (1 μg) and pGEM-4 blue carrier plasmid (14 μg). CAT activities were normalized to β-galactosidase activity from the same extract to correct for differences in transfection efficiency. The relative CAT activity for each construction represents the mean ± S.D. of three separate experiments. The activity produced by RC4CAT/+17 was defined as 1.0.

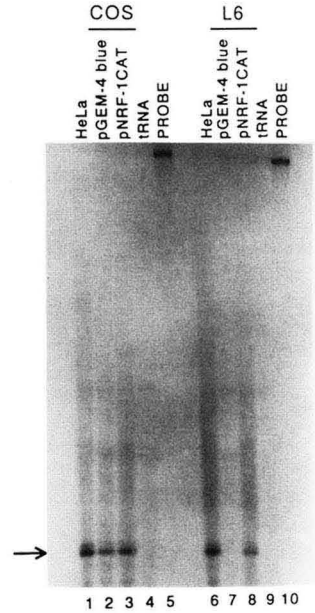

FIG. 5. **Transcription initiation sites of endogenous and transfected NRF-1 genes by ribonuclease protection mapping.** Total RNA was prepared from untransfected HeLa cells (2 μg) (*lanes 1* and *6*) or from COS and L6 cells (10 μg) transfected with either pGEM-4 blue carrier plasmid alone (*lanes 2* and *7*) or the same carrier in combination with pNRF-1CAT (*lanes 3* and *8*). RNA samples were hybridized to an antisense T7 riboprobe spanning the NRF-1 transcription start site, and the hybrids were digested with RNase A and RNase T1 as described under "Experimental Procedures." Yeast tRNA (10 μg) was treated identically as a negative control (*lanes 4* and *9*). *Lanes 5* and *10* contained undigested probe. The position of the protected product is indicated (*arrow*).

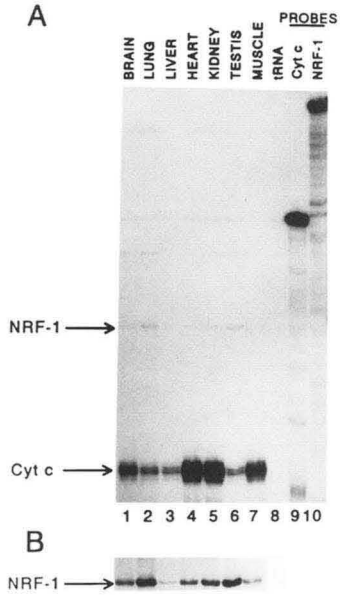

FIG. 6. **Comparison of NRF-1 and cytochrome mRNA levels in rat tissues.** *A*, total RNA (20 μg) prepared by the urea/LiCl method from the tissues indicated above *lanes 1–7* was assayed for cytochrome *c* and NRF-1 mRNAs by RNase protection analysis using specific antisense riboprobes as described under "Experimental Procedures." *Lane 8* contained 20 μg of yeast tRNA treated identically to the tissue RNA samples. *Lanes 9* and *10* contained the undigested cytochrome *c* and NRF-1 riboprobes, respectively. The positions of the RNase-protected NRF-1 (157 bp) and cytochrome *c* (*Cyt c*; 108 bp) products obtained with the riboprobes are indicated (*arrows*). *B*, shown is the extended exposure (3 days) of the same autoradiogram in *A* to facilitate better visualization of the NRF-1-protected product.

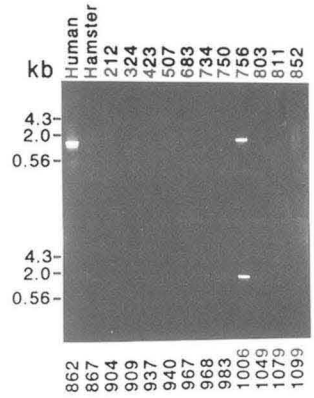

FIG. 7. **PCR analysis of genomic DNA from human-hamster somatic cell hybrids.** Total genomic DNA from human, hamster, or the indicated hybrid cell line was amplified with primers NRF-1 6-2F and NRF-1 6-2R as described under "Experimental Procedures." An aliquot (20 μl) of each PCR was analyzed on a 1% agarose gel containing ethidium bromide and visualized under UV light. A description of the human chromosome content of each of the hybrid cell lines is listed in Table II. Size markers are indicated at the left.

TABLE II
Human chromosome composition of human-hamster somatic cell hybrids

Hybrid cell line	Human chromosome composition[a,b]
212	5 (MD), Y
324	18
423	3
507	3, 5, 12, 14 (65%), 20 (40%), 22 (25%)
683	1 (5–30%), 5 (MD), 12, 14, 19, 21, 22
734	5, 9
750	5, 13, 14, 15, 19
756	5 (MD), 6, 7, 12 (5–30%), 13, 14, 19, 20, 21, Y
803	4 (5–30%), 5, 8, 22, X
811	8, 17, 18
852	2
862	5, 9
867	1, 5, 13, 14, 18, 19
904	5, 6, 12, 16, 21, Y
909	5 (MD), 6, 8, 14, X
937	1, 5, 14, 15, 17 (5–30%), 21
940	5, 20
967	5, 8, 16
968	5, 9, 13
983	5, 10 (15%)
1006	4, 5, 7, 13, 15, 19, 21, Y (5–30%)
1049	5, 11
1079	3, 5
1099	1, 5 (MD), 13, 19, 21, 22

[a] MD indicates the presence of multiple deletions in the particular chromosome.

[b] Numbers in parentheses indicate the average percentage of the specified chromosome present in the particular hybrid line.

NRF-1 transcripts are observed upon extended exposure of the autoradiogram (Fig. 6*B*). The highest levels of NRF-1 mRNA expression are in lung and testis, followed by intermediate levels in kidney, heart, and brain, with the lowest levels in muscle and liver. This contrasts with cytochrome *c* mRNA, which is expressed abundantly in heart and kidney, at an intermediate level in muscle and brain, and at the lowest levels in testis, lung, and liver. Aside from both having their lowest expression in liver, there is little correlation between the expression of cytochrome *c* and NRF-1 mRNAs.

The NRF-1 gene may ultimately prove to be a target for mutations that affect mitochondrial function. To determine its chromosomal location, a panel of human-hamster hybrid DNAs were subjected to PCR amplification using NRF-1 gene-specific primers complementary to segments of introns I-6 and I-8 as indicated in Fig. 1. Electrophoresis of the PCR products on a 1% agarose gel followed by ethidium bromide staining revealed the expected size fragment of 1.5 kb in the lane containing amplification products from human genomic DNA (Fig. 7). The size of the amplified product corresponded to the sum of the sizes of exons CE-5 (158 bp) and CE-6 (197 bp) and the intervening intron I-7 of 1.2 kb (see Fig. 1). No amplification product was observed when hamster DNA was used as the PCR template (Fig. 7), indicating that the primers are human-specific. Analysis of the amplification products obtained using a panel of DNA samples extracted from human-hamster cell hybrids as templates for PCR showed the 1.5-kb fragment only in cell lines 756 and 1006 (Fig. 7). Examination of the human chromosomal composition of the various hybrid lines (Table II) revealed that the only human chromosome present in lines 756 and 1006 and not in the other hybrids is chromosome 7. This result was verified using a second set of human-specific primers derived from the 5′-end of the gene (data not shown).

To confirm the assignment of the NRF-1 locus to chromosome 7 and to further map its position within this chromosome, fluorescence *in situ* hybridization analysis of human metaphase chromosomes was performed. Hybridizations were carried out using biotin-dUTP-labeled ϕ 47-1 DNA and resulted in the specific labeling of the long arm (q) of a group C chromosome, the size and morphology of which were consistent with chromosome 7 (data not shown). To further establish the identity of the labeled chromosome, a chromosome 7 centromere-specific probe was cohybridized to metaphase chromosomes along with labeled ϕ 47-1 DNA from the NRF-1 gene (see Fig. 1). In each of the four separate determinations shown in Fig. 8, hybridizing bands were detectable at the centromere and at a position distal to the centromere in the chromosome 7 pair. Measurements performed on 10 specifically labeled chromosomes 7 indicated the NRF-1 locus to be present at a position that corresponded to 59% of the distance

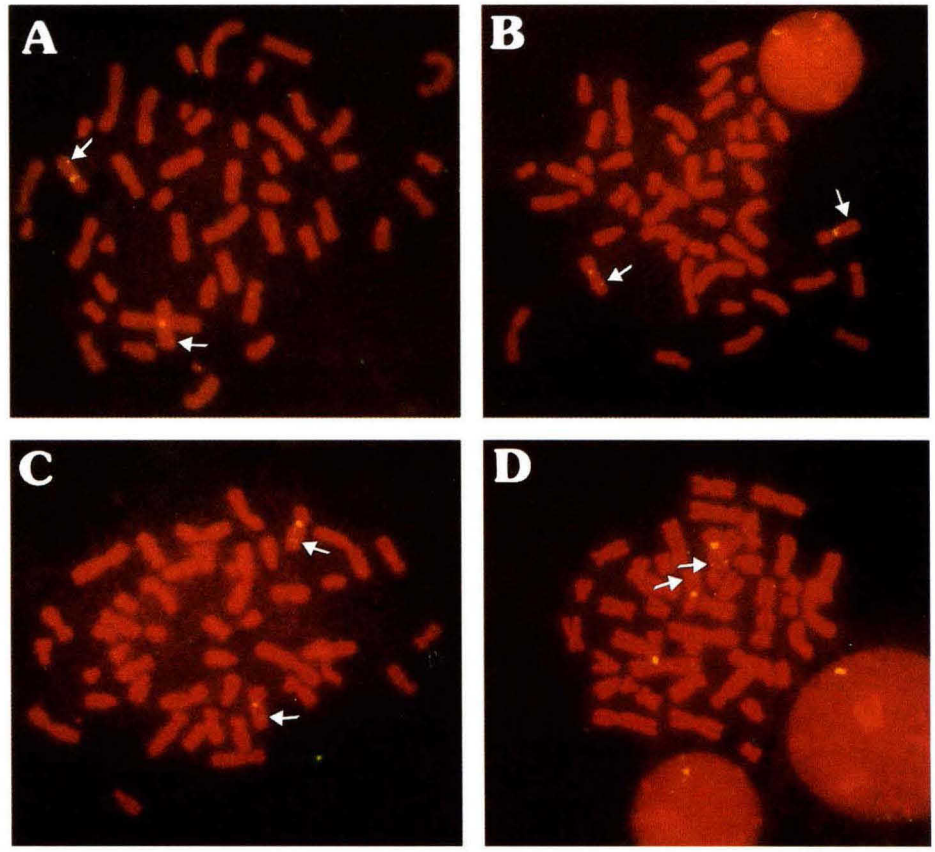

FIG. 8. **Fluorescent *in situ* hybridization analysis of human metaphase chromosomes with a NRF-1 genomic probe.** *A–D*, *in situ* hybridization of biotin-labeled ϕ 47-1 DNA to four separate sets of human metaphase chromosomes derived from phytohemagglutinin-stimulated peripheral blood lymphocytes. Signals were detected by incubation of the post-hybridization slides in fluorescein-conjugated avidin. The specific labeling of the chromosomes at 7q31 is indicated by the *arrows*. The chromosome 7 homologues are identified by centromere-specific staining.

from the centromere to the telomere of 7q. This position corresponds to band 7q31.

DISCUSSION

The results presented here are consistent with a single-copy gene encoding NRF-1 in humans. Its genomic organization consists of 11 exons spread over 65 kb, and the polypeptide specified by the coding exons is identical to that predicted from the NRF-1 cDNA (11). The mRNA 5′-terminus cloned by RACE-PCR contains the two 5′-terminal exons, and genomic clones containing these exons are overlapping and contiguous with the remainder of the gene. An active promoter residing in a genomic fragment containing the 5′-terminal exon initiates transcription at the same sites utilized by the endogenous gene and therefore represents a functional promoter *in vivo*. These results establish that the gene described here is responsible for the expression of human NRF-1.

There is no evidence that the NRF-1 gene is a member of a transcription factor family of related genes. Despite a striking degree of sequence conservation within the NRF-1 DNA-binding domains between humans and lower eukaryotes (11), no related human family members were detected by genomic hybridization analysis. All of the recombinant phage clones obtained represented overlapping segments of a single genomic locus. The restriction fragments from these genomic clones that hybridized to the NRF-1 cDNA probe accounted for all of those detected in total genomic DNA. Finally, only a single NRF-1 chromosomal locus mapping to 7q31 was detected by fluorescence *in situ* hybridization analysis using a probe encompassing a highly conserved region of the DNA-binding domain (Fig. 1, *CE-2* and *CE-3*).

The presence of NRF-1 recognition sites in many nuclear genes whose products function in the mitochondria suggests a role for NRF-1 in nuclear-mitochondrial interactions (10, 11, 13). In at least two of these genes, mtTFA and 5-aminolevulinate synthase, the NRF-1 sites have been found to be important determinants of promoter function (13, 14). Transcriptional activation of D-loop promoters by mtTFA facilitates the synthesis of mitochondrial RNAs as well as heavy strand replication primers (1, 4). Thus, NRF-1 may help coordinate the synthesis of respiratory subunits from both genomes and contribute to the maintenance of mtDNA through its effects on mtTFA expression.

Recently, mtTFA levels were found to be reduced in patients with mtDNA depletion (17, 18). Elevated mtTFA levels have also been found in the ragged-red muscle fibers of patients with mtDNA deletions that lead to the abnormal proliferation of mtDNA (17). Thus, mtTFA seems to vary with mtDNA in these individuals. One possibility is that the reduction in mtTFA is not a primary cause of mtDNA depletion, but rather a consequence of reduced mtDNA levels. The fact that a patient with reduced mtDNA and mtTFA also had low mtTFA RNA levels argues that the primary defect may be in the synthesis or degradation of mtTFA RNA (18). Cell lines that are transiently depleted of mtDNA with dideoxycytidine (32) had reduced mtTFA but increased mtTFA RNA, suggesting that expression is up-regulated at a pre-translational level in response to the loss of mtDNA (18).

One intriguing possibility is that mutation in a regulatory gene encoding NRF-1 or some other factor may lead to reduced expression of mtTFA and consequently the loss of mtDNA. Similarly, reduction of mtDNA by dideoxycytidine treatment or impaired respiratory function from mtDNA mutation may lead to the activation of mtTFA transcription by NRF-1. However, the fact that mtTFA RNA levels are unchanged in established cell lines lacking mtDNA as a result of propagation in ethidium bromide (17, 33) is inconsistent with such a feedback mechanism and with the results obtained with dideoxycytidine (18). It will be of considerable interest to determine whether mutations

in the NRF-1 gene can lead to aberrant mtDNA copy number control and to impaired oxidative metabolism. The isolation, characterization, and chromosomal assignment of the human NRF-1 gene should aid in evaluating its potential role in mtDNA depletion and possibly other nuclear gene defects resulting in mitochondrial disease.

Acknowledgments—We thank Dr. J. K. Reddy for the human placenta genomic library and Dr. C. A. Bradfield for the human-hamster somatic cell hybrid DNA panel used in the PCR experiments.

REFERENCES

1. Clayton, D. A. (1992) *Int. Rev. Cytol.* **141,** 217–232
2. Wallace, D. C. (1992) *Annu. Rev. Biochem.* **61,** 1175–1212
3. Kruse, B., Narasimhan, N., and Attardi, G. (1989) *Cell* **58,** 391–397
4. Shadel, G. S., and Clayton, D. A. (1993) *J. Biol. Chem.* **268,** 16083–16086
5. Parisi, M. A., Xu, B., and Clayton, D. A. (1993) *Mol. Cell. Biol.* **13,** 1951–1961
6. Evans, M. J., and Scarpulla, R. C. (1989) *J. Biol. Chem.* **264,** 14361–14368
7. Gopalakrishnan, L., and Scarpulla, R. C. (1994) *J. Biol. Chem.* **269,** 105–113
8. Virbasius, J. V., and Scarpulla, R. C. (1991) *Mol. Cell. Biol.* **11,** 5631–5638
9. Virbasius, J. V., Virbasius, C. A., and Scarpulla, R. C. (1993) *Genes & Dev.* **7,** 380–392
10. Evans, M. J., and Scarpulla, R. C. (1990) *Genes & Dev.* **4,** 1023–1034
11. Virbasius, C. A., Virbasius, J. V., and Scarpulla, R. C. (1993) *Genes & Dev.* **7,** 2431–2445
12. Gugneja, S., Virbasius, J. V., and Scarpulla, R. C. (1995) *Mol. Cell. Biol.* **15,** 102–111
13. Virbasius, J. V., and Scarpulla, R. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 1309–1313
14. Braidotti, G., Borthwick, I. A., and May, B. K. (1993) *J. Biol. Chem.* **268,** 1109–1117
15. Moraes, C. T., Shanske, S., Tritschler, H.-J., Aprille, J. R., Andreetta, F., Bonilla, E., Schon, E. A., and DiMauro, S. (1991) *Am. J. Hum. Genet.* **48,** 492–501
16. Tritschler, H.-J., Andreetta, F., Moraes, C. T., Bonilla, E., Arnaudo, E., Danon, M. J., Glass, S., Zelaya, B. M., Vamos, E., Telerman-Toppet, N., Shanske, S., Kadenbach, B., DiMauro, S., and Schon, E. A. (1992) *Neurology* **42,** 209–217
17. Larsson, N.-G., Oldfors, A., Holme, E., and Clayton, D. A. (1994) *Biochem. Biophys. Res. Commun.* **200,** 1374–1381
18. Poulton, J., Morten, K., Freeman-Emmerson, C., Potter, C., Sewry, C., Dubowitz, V., Kidd, H., Stephenson, J., Whitehouse, W., Hansen, F. J., Parisi, M., and Brown, G. (1994) *Hum. Mol. Genet.* **3,** 1763–1769
19. Feinberg, A. P., and Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13
20. Scarpulla, R. C., Agne, K. M., and Wu, R. (1981) *J. Biol. Chem.* **256,** 6480–6486
21. Kaslow, D. C. (1986) *Nucleic Acids Res.* **14,** 6767
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
23. Scarpulla, R. C., Agne, K. M., and Wu, R. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 739–743
24. Burke, J. F. (1984) *Gene (Amst.)* **30,** 63–68
25. Scarpulla, R. C., and Wu, R. (1983) *Cell* **32,** 473–482
26. MacGregor, G. R., Mogg, A. E., Burke, J. F., and Caskey, C. T. (1987) *Somatic Cell Mol. Genet.* **13,** 253–265
27. Evans, M. J., and Scarpulla, R. C. (1988) *Mol. Cell. Biol.* **8,** 35–41
28. Rosenthal, N. (1987) *Methods Enzymol.* **152,** 704–720
29. Virbasius, J. V., and Scarpulla, R. C. (1988) *J. Biol. Chem.* **263,** 6791–6796
30. Krieg, P. A., and Melton, D. A. (1987) *Methods Enzymol.* **155,** 397–415
31. Virbasius, J. V., and Scarpulla, R. C. (1990) *Nucleic Acids Res.* **18,** 6581–6586
32. Chen, C.-H., and Cheng, Y.-C. (1989) *J. Biol. Chem.* **264,** 11934–11937
33. Tominaga, K., Hayashi, J., Kagawa, Y., and Ohta, S. (1993) *Biochem. Biophys. Res. Commun.* **194,** 544–551