# Human gene for the RNA polymerase II seventh subunit (hsRPB7): structure, expression and chromosomal localization

Timothy J. Schoen [a], Settara C. Chandrasekharappa [b], Siradanahalli C. Guru [b], Krzysztof Mazuruk [a], Gerald J. Chader [a], Ignacio R. Rodriguez [a, *]

[a] *National Eye Institute, National Institutes of Health, 6 Center Drive MSC2740, Bldg. 6, Rm. 304, Bethesda, MD 20892, USA*
[b] *The National Center for Human Genome Research, National Institutes of Health, Bethesda, MD 20892, USA*

## Abstract

The human gene for the seventh largest subunit of RNA polymerase II complex, hsRPB7 was cloned, sequenced and mapped. This complex is an integral part of the transcription-coupled DNA repair mechanism and has been shown to be involved in several human genetic diseases and implicated in many others. The hsRPB7 gene consists of 8 exons and spans approximately 5.1 kb. Southern blots of genomic and cloned DNA suggest that hsRPB7 is coded for by a single gene. Using human radiation hybrids and YACs, the gene was localized to 11q13.1, within 70 kb of marker D11S1765. The sequence of the 5' flanking region does not contain a TATA element, but does contain several Sp1 binding sites, an AP-1 site and a novel inverted polymorphic GATA tandem repeat. This novel GATA repeat can be used for linkage analysis. The hsRPB7 gene seems to be highly conserved among eukaryotic species, showing general sequence conservation to yeast and *Drosophila*. Northern blot analysis reveals a high degree of tissue-specific expression. For example, adult retina, brain and kidney exhibit a relatively high level of expression. A moderate level of expression is observed in heart, lung, testis, cornea, retinal pigmented epithelium/choroid and placenta with a lower level of expression in the uterus, small intestine and skeletal muscle. A very low level of expression was observed in stomach and liver. Comparison between four fetal and adult tissues also demonstrate a surprising level of developmental specificity. Expression in fetal retina is considerably lower than fetal brain but similar to adult retina. © 1997 Published by Elsevier Science B.V.

*Keywords:* Chromosome 11; Mapping; GATA repeat

## 1. Introduction

RNA polymerase II (Pol II) is the enzymatic complex responsible for transcription of genes that result in messenger RNA (mRNA) production. RNA polymerase II is composed of 10–14 subunits ranging in size from 220 to 10 kDa [1]. This complex interacts with the promoter regions of genes as well as with a variety of elements and transcription factors to determine essentially all of the parameters that govern transcription (e.g., tissue and developmental specificity, stress response, etc., see review [2]). Presently, human cDNAs have been isolated and chromosomal localization performed on a total of seven of the Pol II complex genes [3–5]. However, to date, the com-

* Corresponding author. Fax: +1 (301) 4020750. E-mail: ignacio@helix.nih.gov

plete genomic sequence has been reported for only the 220 kDa [4] and 14.5 kDa [5] subunits.

During the process of screening a macula-enriched subtraction library [6], we isolated a cDNA clone that was determined, during a BLAST [7] database search, to be hsRPB7 [8]. Although hsRPB7 message expression is only two-fold enriched between macula and peripheral retina, its chromosomal localization to 11q13.1 made this gene a possible candidate for several genetic diseases located in this area. Due to the overall importance of the Pol II complex and the possible involvement of this gene in genetic disease, we have characterized and localized the human gene for hsRPB7 as well as determined its relative level of expression in a variety of human tissues.

## 2. Materials and methods

### 2.1. 5′ Rapid amplification of cDNA ends (5′ RACE)

The 5′ RACE was performed using a previously published technique [9]. Two hsRPB7 specific primers were used: 5253 ($^{5'}$TGTGCCACTTCGTCT-TCGAGA$^{3'}$) and 5440 ($^{5'}$TACAGAACGAAG-TAGAGAGCTG$^{3'}$).

### 2.2. Screening of a human genomic library

A human genomic library constructed in PWE-15 (Clontech, Palo Alto, CA) was screened using a $^{32}$P-labelled PCR product from the hsRPB7 cDNA. The PCR product was labeled with [$^{32}$P]dCTP to a specific activity of approximately $5 \times 10^9$ cpm/$\mu$g using a PCR-labeling technique [10]. Approximately 300 000 clones were screened resulting in one positive clone pTS14 of approximately 30 kb in length. A human YAC and a PAC library was screened by Genome Systems Inc. (St. Louis, MO). Two YAC clones, 737-H7 and 782-G7 and one PAC clone p11929 were identified.

### 2.3. DNA sequencing

The hsRPB7 gene was sequenced directly from the cosmid clone pTS14 as well as from PCR products and subclones derived from pTS14. Fluorescent se-

quencing was performed in a Perkin-Elmer/Applied Biosystems (ABI) model 370A instrument. The sequencing was performed using ABI PRISM™ Dye Terminator cycle sequencing kit following the manufacturer's protocol. In general, 0.5 pmol of template and 3 pmol of primer were used per sequencing reaction. All other details are provided in the ABI's manual included in the sequencing kit. The reactions were purified using Select-D G-50 columns purchased from 5Prime-3Prime (Boulder, CO).

### 2.4. Chromosomal sub-localization by PCR

The human monochromosomal somatic cell hybrid panel was purchased from NIGMS Human Genetic Mutant Cell Repository at the Coriel Institute for Medical Research (Camden, NJ). PCR amplification was performed using the hsRPB7 primers 4823, $^{5'}$GCTACAAGAAGCACCAGATGC$^{3'}$ and 4824,
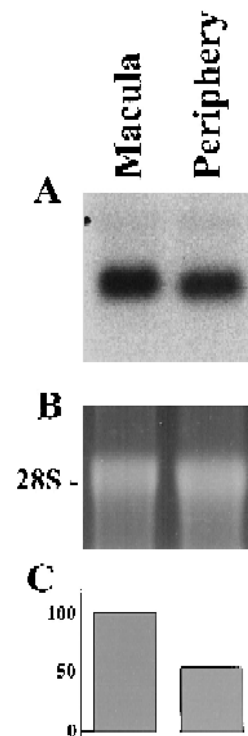


Fig. 1. Expression of hsRPB7 in macula vs. peripheral retina. Northern blot analysis of total RNA (5 $\mu$g) from monkey macula and peripheral retina. The blot was probed with hsRPB7 cDNA as described in Section 2. Panel A shows the autoradiogram after 8 h of exposure. Panel B shows the ethidium bromide staining of the 28S band. The relative expression normalized to the 28S ribosomal RNA band is shown in the graph in panel C.

$5'$TACTTGGGGCTTGTAAGCTGA$3'$. The PCR reaction mixture contained approximately 300 ng of genomic DNA, 25 pmol of each primer, 200 $\mu$M dNTP's, 10 mM Tris (pH 8.8), 50 mM KCl, 1.5 mM MgCl$_2$ and 2.5 units of Taq polymerase (Stratagene, La Jolla, CA). Amplification was performed in one cycle of 2 min at 95°C, followed by 30 cycles of a 45-s denaturation step at 95°C, a 1-min annealing step at 55°C and a 1-min extension at 72°C, with a final extension step of 5 min. The YAC clones were tested by PCR with primers specific for several markers in the area. The microsatellite marker D11S1765 ($5'$CAGAAATGCCACCCAGAGAG$3'$, $5'$TTCCG-GAGTTTGCACAATCT$3'$) was found to co-localize with hsRPB7 in YACs 782-G7.

The hsRPB7 primers and the PCR conditions as described above were used to screen the following somatic cell hybrids containing parts of the chromosome 11: J1-11, J1-44, J1-46, R28-4D, and G35F1A. In addition, the hybrid J1 with entire chromosome 11 was also used along with hamster mouse and human DNA controls. All the somatic cell hybrids were obtained from Dr. Carol Jones (J1, J1-11, J1-44, J1-46, R28-4D) except G35F1A which was obtained from Dr. Philippe Couillin. The region of chromosome 11 present in the somatic cell hybrids was described earlier [11,12]. The hybrid G35F1A was described in Ref. [13].

## 2.5. Southern blot analysis of human genomic DNA

Samples containing 10 $\mu$g of high molecular weight human genomic DNA (Clontech) and 300 ng of cosmid DNA were digested with selected restriction enzymes for 16 h at 30°C. The fragments were separated on a 0.7% (w/v) agarose (0.5 × TBE) gel at 1 V/cm, stained in ethidium bromide (0.5 g/ml), photographed and transferred to a nylon membrane as described for Northern blots. Following the transfer, the blot was prehybridized in a solution containing 6 × SSC, 10 × Denhardt's and 1.0% SDS for 1 h at 42°C. Hybridization was performed at 42°C for 18 to 24 h in a solution of 50% formamide, 5 × SSC, 1 × Denhardt's, 1.0% SDS and 10% dextran sulfate and approximately 10$^6$ cpm/ml of the $^{32}$P-labeled full-length hsRPB7 cDNA PCR product. Following hybridization, the blots were washed in 2 × SSC containing 1.0% SDS for 15 min at room temperature

and transferred to a solution containing 0.2 × SSC with 0.1% SDS at 65°C for 2 × 15 min. Autoradiography was performed using X-Omat™ film with an intensifying screen at −80°C.

## 2.6. Northern hybridization

Total RNA was isolated from embryonic and adult retina, cornea and retinal pigmented epithelium(RPE)/choroid using the RNAazol method (Cinna/Biotecx Laboratory, Friendswood, TX). Human total RNA for all other tissues was obtained from Clontech (Palo Alto, CA). Samples containing total RNA (5 $\mu$g) were electrophoresed on a 1% agarose formaldehyde gels at 25 volts for 3 h. Following electrophoresis, the gels were photographed and soaked in 20 × SSC for 15 min to remove excess formaldehyde and then blotted to nylon membranes (Schleicher and Schuell, Keene, NH) using a Posiblot apparatus (Stratagene). Following the transfer, the



Fig. 2. Sequence of the human hsRPB7 cDNA. The complete sequence of the human hsRPB7 cDNA was put together from a combination of our initial isolate, the GenBank file accession no. U20659 [8], genomic sequence and 5′ RACE. Sequence obtained from 5′ RACE is shown in bold and the primers used in the 5′ RACE amplification shown in arrows. The complete sequence is 826 bp in length. An open reading frame of 172 amino acids is shown above the nucleotide sequence and the position of the exon/intron junctions are also indicated. An imperfect polyadenylation sequence is shown in double underline.
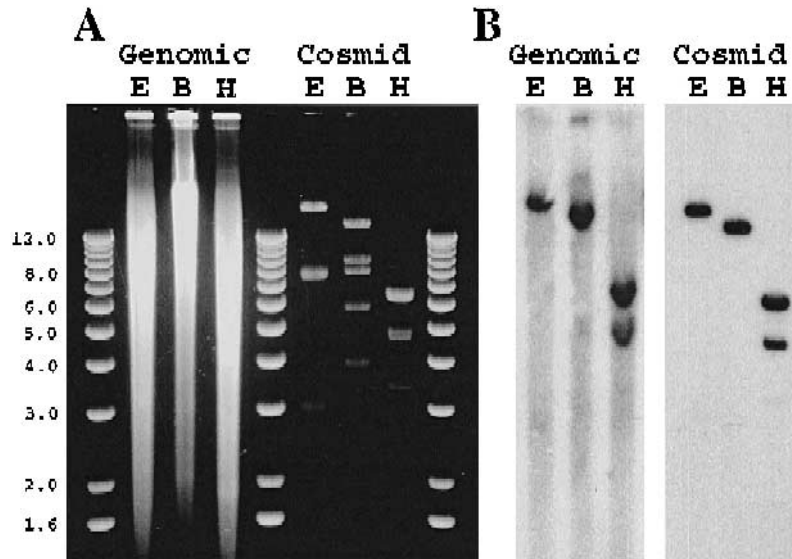
Fig. 3. Southern blot hybridization of human genomic and cosmid DNA. Human genomic DNA (10 $\mu$g) and the hsRPB7 genomic cosmid clone (pTS14, 300 ng) was digested with *Eco*RI (E), *Bam*HI (B), or *Hin*dIII (H). The DNA was separated in a 1% agarose TAE gel and visualized with ethidium bromide. Panel A shows the gel before blotting and panel B shows the blot after hybridization with a ^{32}P-labeled full-length cDNA probe generated by PCR.

blot was washed for 5 min in $20 \times$ SSC, blotted damp dry and cross-linked using a UV Stratalinker (Stratagene). Blots were probed as described above and autoradiography was performed using X-Omat™ film with an intensifying screen at $-75°$C. An estimation of the hybridizing band size was obtained by comparison with the migration of RNA molecular weight standards (Life Technologies, Bethesda, MD). To normalize for variations in loading, the ethidium-stained 28S and 18S bands in the gel were photo-

graphed and the gel negative along with the hsRPB7 autoradiograms were scanned using an Ultroscan XL Laser Densitometer with Gel Scan software (Pharmacia, Uppsala, Sweden).

### 2.7. 'Zoo' Southern blot

Samples containing 8 $\mu$g of high molecular weight genomic DNA from human, monkey, pig, dog, cow,

Table 1
Intron-exon junctions in the human hsRPB7 gene

| Exon no. | Size (bp) | Exon | Intron | Size (bp) | Intron | Exon | Exon no. |
|---|---|---|---|---|---|---|---|
| | | | | | 5′ flanking | TAGTCGCAC | 1 |
| 1 | 118 | TTCTACCAT | gtgagcagg... | 133 bp | ...tctccgcag | ATCTCCCTA | 2 |
| 2 | 110 | CACAGGGAA | gtgagtgtc... | 965 bp | ...tccttgcag | GTATGGCTT | 3 |
| 3 | 160 | GTCAACAAG | gtgagacca... | 2157 bp | ...tttttctag | GTTGGACTC | 4 |
| 4 | 51 | TCTCGACAT | gtaagtctg... | 107 bp | ...tcctttcag | TCCATCCCT | 5 |
| 5 | 66 | ATGGATGAG | gtgagtgga... | 247 bp | ...tctcctcag | GATATTGTG | 6 |
| 6 | 72 | AATGACATT | gtgagtctt... | 518 bp | ...ttcttgcag | TTTGCTATT | 7 |
| 7 | 34 | ATTACTTGG | gtgagtgcc... | 217 bp | ...gtttcgcag | GGCTTGTAA | 8 |
| 8 | 217 | TTGGTTCTC | atggaagtg... 3′ flanking | | | | |

The complete sequence of the human hsRPB7 gene can be obtained from GenBank accession no. U52427. This table shows the first and last nine base pairs of each intron and exon starting with exon 1. The sizes (i.e., base pairs) given for the lengths of the introns include the bases shown.
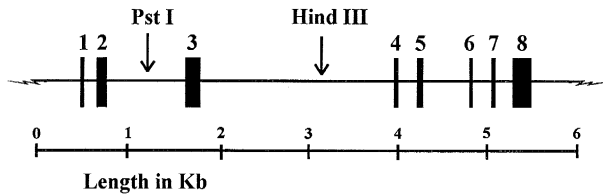
Fig. 4. Scale map of the human hsRPB7 gene. Exons are indicated by solid boxes. Positions of the *Hin*dIII and *Pst*I sites used in subcloning are shown.

mouse, rabbit, rat, chicken, yeast and *Drosophila* (Clontech, Palo Alto, CA) were restriction digested with *Eco*RI, separated on a 0.7% agarose gel, blotted to a nylon membrane and probed with the full-length hsRPB7 cDNA PCR product. Blots were washed at high stringency as described above

## 3. Results

### 3.1. Isolation of hsRPB7 cDNA and 5′ RACE

The hsRPB7 cDNA was isolated during a solid-phase subtraction between the macula and peripheral areas of the retina as previously described [6]. Northern blot analysis of total RNA from monkey macula and peripheral retina demonstrates that hsRPB7 is enriched about 2-fold in the macula versus peripheral retina (Fig. 1). The original isolate was identified by performing a GenBank search using the National Center for Biological Information BLAST server program [7]. The search identified this clone as hsRPB7, reported by Khazak et al. [8]. Further analysis of this clone using a solid-phase 5′ RACE technique [9], revealed an additional 78 bp of 5′ untranslated se-
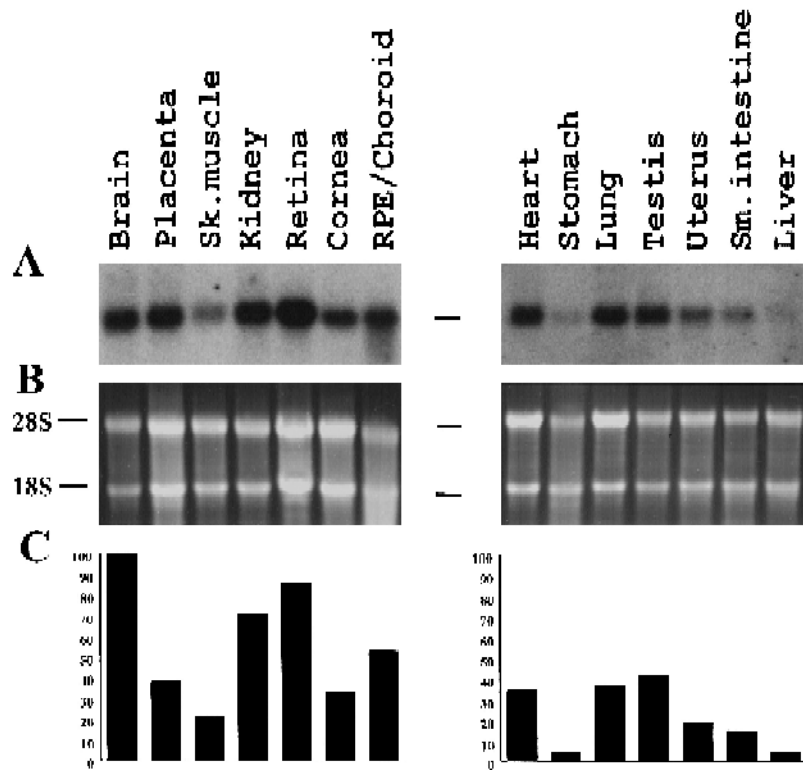


Fig. 5. Multi-tissue expression analysis of the human hsRPB7 mRNA. Northern blot analysis was performed on total RNA $(5 \ \mu g)$ from different adult human tissues. The autoradiograms of the blots probe with hsRPB7 cDNA are shown in panel A and the ethidium bromide stained gels are shown in panel B. The autoradiograms were scanned densitometrically and normalized to the relative intensity of the 28S band. The relative expression, normalized for loading are shown in the graph in panel C.

quence not reported by Khazak et al. [8], bringing the full-length hsRPB7 cDNA to 826 bp (Fig. 2).

### 3.2. Isolation of the hsRPB7 gene

A human genomic cosmid library (Clontech) was screened using an hsRPB7 PCR product obtained from human retina cDNA and a single cosmid clone (pTS14) was isolated. Before proceeding to subclone and sequence the hsRPB7 gene, a Southern blot analysis was performed on *Bam*HI, *Eco*RI and *Hin*dIII digested human genomic and cosmid DNA. This experiment was performed to see if the clone contained the entire gene and to determine if there were any hsRPB7 pseudogenes present in genomic DNA. The results of this experiment can be seen in Fig. 3. Comparison between hybridization patterns of our cosmid isolate pTS14 and the genomic DNA indicate that the entire hsRPB7 gene is contained in cosmid pTS14 and that hsRPB7 is a single gene.

The hsRPB7 gene was completely sequenced using a combination of direct sequencing of the cosmid, cosmid-derived subclones and PCR products. The entire sequence for this gene can be obtained from GenBank accession no. U52427. The hsRPB7 gene consists of 8 exons and spans approximately 5 kb. The intron/exon junctions were determined by direct sequencing of the cosmid using intron- and exon-specific primers. The sequence of the intron/exon junctions, along with the size of the corresponding introns and exons is shown in Table 1. All splice junctions conform to the GT/AG rule. A scaled map of the gene is shown in Fig. 4.

### 3.3. Multi-tissue Northern blot

In order to determine relative expression of hsRPB7 mRNA in adult human tissues, Northern blot analysis was performed (Fig. 5A). The autoradiographs were scanned and normalized to the relative intensity of the 28S ribosomal RNA band stained with ethidium bromide (Fig. 5B) and the results are graphically represented in Fig. 5C. Surprisingly, a wide range of expression was observed. A high level of expression was detected in brain, retina and kidney, with lower expression in the retinal pigmented epithelium/ choroid, placenta, lung, heart, cornea and skeletal
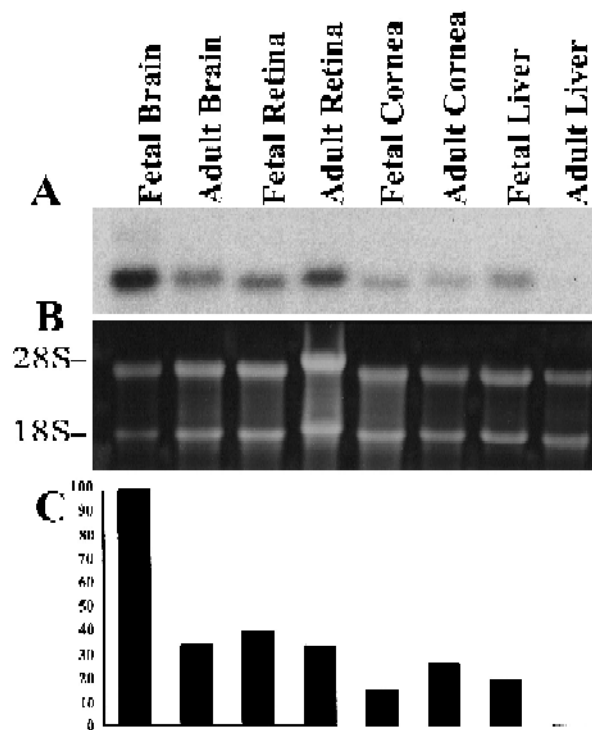


Fig. 6. Developmental difference in expression of hsRPB7 mRNA. Total RNA (5 $\mu$g) from fetal (10–14 weeks) and adult tissues were separated side by side. Panel A shows the autoradiogram and panel B the ethidium bromide staining. The autoradiograms were scanned densitometrically and normalized to the relative intensity of the 28S band. The relative expression, normalized for loading are shown in the graph in panel C.

muscle. Interestingly, hsRPB7 mRNA is barely detectable in adult stomach and liver.

### 3.4. Developmental Northern blot

Because of the variable expression of hsRPB7 in adult human tissues, it was of interest to ascertain whether or not hsRPB7 expression varied during development. Therefore, samples of RNA from fetal (10–14 weeks) and adult brain, retina, cornea and liver were subjected to Northern blot analysis (Fig. 6A). As explained above, hsRPB7 autoradiographs were normalized to the relative intensity of the 28S band stained with ethidium bromide (Fig. 6B) and the results are graphically represented in Fig. 6C. The highest expression of hsRPB7 is observed in fetal brain, over a two-fold higher expression than that

observed in adult brain. A similar pattern of expression is observed with liver where hsRPB7 is high during development and almost non-detectable in the adult. hsRPB7 mRNA expression in the fetal retina is only slightly higher than that of the adult retina. In contrast to the brain, retina and liver, the expression of hsRPB7 mRNA is actually lower in the fetal cornea as compared to adult cornea.

### 3.5. Chromosomal sub-localization

Chromosomal localization by PCR using a human–rodent somatic cell hybrid panel revealed that hsRPB7 is located on chromosome 11 (data not shown). Since several degenerative retinal diseases are located on chromosome 11, it was of interest to determine the precise sub-localization of the hsRPB7
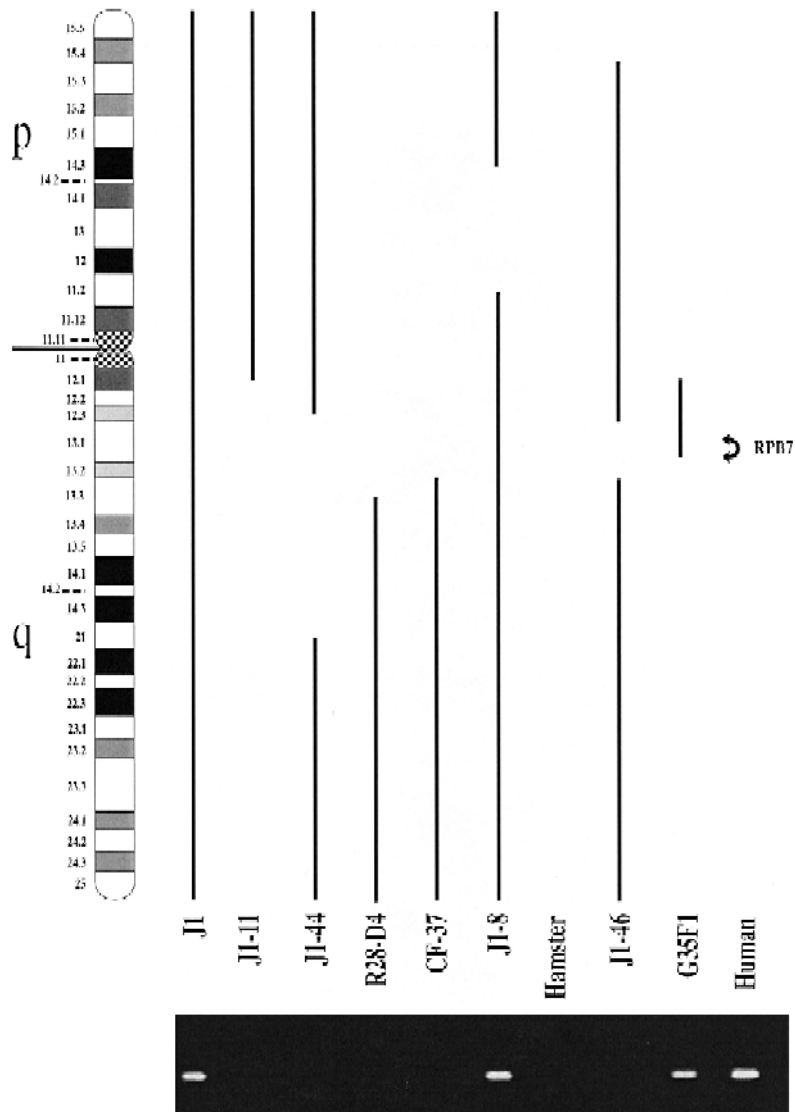


Fig. 7. Sublocalization of hsRPB7 on chromosome 11 using a radiation-hybrid panel. DNA (200 ng) from several radiation hybrids were used as template to amplify hsRPB7 gene using oligos 4823/4824. The chromosome 11 idiogram and the cell lines are listed to show the approximate location of the hsRPB7 gene.
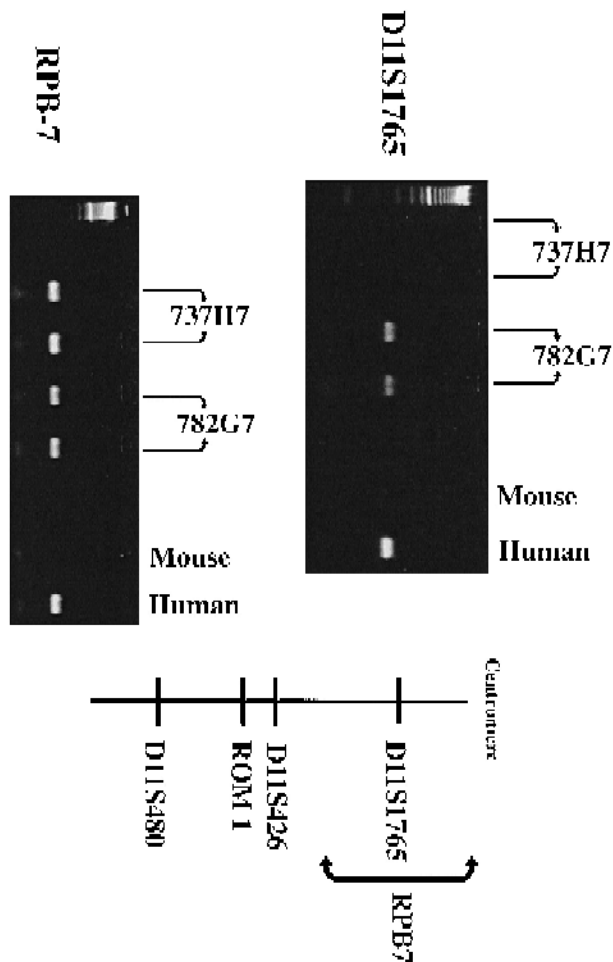
Fig. 8. Association of hsRPB7 to marker D11S1765 using YACs. YACs 737H7 and 782G7 which contain the hsRPB7 gene were tested with several markers in the region. The DNA from the corresponding YACs was used as template to amplify the hsRPB7 and D11S1765 (see Section 2 for primer sequence and PCR conditions).

gene. Fig. 7 shows the results of PCR analysis of several radiation hybrids containing portions of chromosome 11 [11–13]. The finding that G35F1A yields a positive PCR product while J1-46 is negative indicates that the hsRPB7 gene is located between 11q13.1 and 11q12.2. To further define the position of the hsRPB7 gene, PCR was performed using DNA from two YAC clones located in this region (Fig. 8). The finding that YAC 782G7, which is only 70 kb in length, yields a positive PCR product with the D11S1765 primers as well as with the hsRPB7 primers, suggests that hsRPB7 gene is located within

70 kb of D11S1765. A PAC (p11929) clone containing the hsRPB7 gene was found to be negative for D11S1765.

### 3.6. Analysis of 5′ flanking sequence

Analysis of 5′ flanking sequence failed to uncover any TATA or CAAT putative promoter elements, but did reveal several putative Sp1 binding sites at positions −4, −29, −96 and −109, and a single AP-1 binding site at position −286 (Fig. 9). In addition, an inverted polymorphic $(GATA)_{18}$ tetranucleotide repeat is present starting at positions −594 to −520. The oligonucleotides $^{5'}$CAAGCGCTTAGCA-CAGTGTCT$^{3'}$ and $^{5'}$AAGAGCGAAACTCCATCT-CAACAG$^{3'}$ will amplify a 247 bp product from our genomic clone and will generate two products of different size from most genomic DNA samples (data not shown). Our preliminary results indicate that this GATA repeat will be suitable for linkage analysis.

### 3.7. Species conservation of the hsRPB7 gene

In order to determine the degree of conservation of the hsRPB7 gene, a Southern blot of EcoRI-digested

```
-750   agtgatagcatttatttatttatttattttatttattgagatggagtttc
-650   gctcttgttgcccaggcatctaggaagaaagcagtaagaggtagcattta
-600   tcaatctatctatctatctatctatctatctactatctatctatctatct
-550   atctatctatctatctatctatctatctatctgttgagatggagtttcgc
-500   tcttgttgcccaggctggagttcagtggcaagatctcgggccgcctccca
-450   ggttcaagcaattcagatagttgccgaccgggcgcggtggctcactcctg
-400   taatcccaacactttgggaggccgaggtgggtggatcacctgaggttggg
-350   agtttgagaccagcctgaccaacatggagaaaccccatctctactaaaaa
-300   tacaaaattagtcaggtgtggtggcacatgcctttaatcccagctactcg
-250   ggaggatgaggcaggagaatcgcttgaacccgggaggcggaggttgcagt
-200   gagccgagatcgcgccactgcactccagcctgggcaacagagggagactc
-150   cgcctcaaaaaaaaaagttgtcgtttgatggctgggggcgggggtgagg
-100   ggcgggcttagtggttctcgcccacgacggcgaggcccaatcagacccg
-50    cggagctggtggagttccgccctcgcggaggaggtggtgttccgccctcc
+1     TAGTCGCAGCCAAGCGCGGAACTGGGGTTGCGCGTCTAAGTGTTTCCGGT
+50    GGATTCCCAGGGACTGTCGGAGGTGTGGACTCTGCCTGCCTACCTGGTCT
+100   GGGAAGATGTTCTACCATgtgagcagg
```

Fig. 9. Analysis of promoter region of hsRPB7 gene and polymorphic GATA. The sequence of the hsRPB-7 gene from the transcription starting position −1 to −750 is shown in lower case. Exon 1 is shown in upper case letters. A polymorphic inverted GATA repeat from −594 to −520 is underlined. Sp1 binding sites at positions −4, −29, −96 and −109 and a single AP-1 site at position −286 are shown in bold.
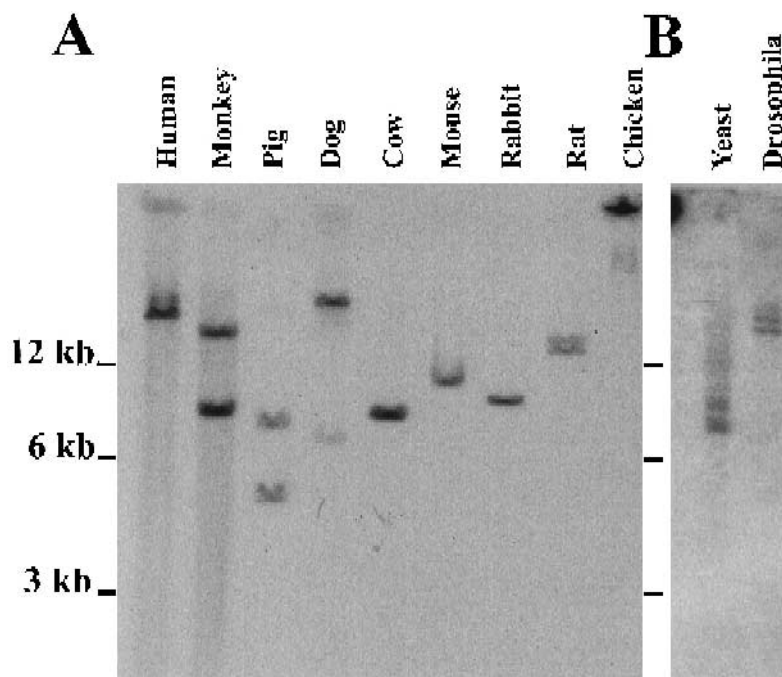
Fig. 10. Southern 'Zoo-blot' containing *Eco*RI-digested genomic DNA from eukaryotic species. Southern 'Zoo-blot' containing *Eco*RI-digested genomic DNA from 9 eukaryotic species. The digested DNA (8 $\mu$g) was separated on a 0.7% agarose gel and blotted. The blot was probed with a human hsRPB7 cDNA probe. After hybridization, the blot was washed under stringent conditions, $0.2 \times$ SSC, 0.1% SDS, twice for 30 min at 65°C and the autoradiographic exposure time was 24 h for all species except for the yeast and *Drosophila*, which were exposed for 72 h.

genomic DNA from several different species was probed with a full-length hsRPB7 cDNA probe under stringent conditions. The results demonstrate that the hsRPB7 gene is highly conserved among vertebrates (Fig. 10). On prolonged exposure, the human hsRPB7 also seems to recognize specific bands in yeast and *Drosophila* DNA. The genomic blot containing the vertebrate DNA samples was exposed for 24 h, whereas the yeast and *Drosophila* DNA was exposed for 72 h.

## 4. Discussion

Our interest in cloning the hsRPB7 gene was 3-fold: its chromosomal localization, its high and differential expression in the macula region of the retina and its fundamental importance in RNA transcription. Based on our radiation hybrid and YAC analysis, the hsRPB7 gene is located in chromosome 11 band q13.1, closely associated with the marker D11S1765. Attempts to associate this gene with other

markers in the area using our YACs and PAC have been unsuccessful so far. However, we have localized a new polymorphic GATA repeat in the 5′ flanking region of hsRPB7 which can be used for linkage analysis (Fig. 9).

The higher expression of the hsRPB7 mRNA in the macula than in peripheral retina (Fig. 1) is of special interest to us due to the critical importance of the macula in a variety of important diseases of the eye. This observation coupled with its chromosomal localization makes hsRPB7 a possible candidate for Best's macular dystrophy, a disease preferentially affecting the macula area of the retina which impairs central vision [14]. Although hsRPB7 mRNA is highly expressed in a variety of tissues besides the retina, this does not necessarily exclude it from causing a retinal degeneration. Mutations in a number of genes that are generally expressed have been found clinically to affect only the retina [15,16].

The wide-ranging expression of hsRPB7 in different tissues has been shown previously by Khazak et al. [8]. The reason(s) for these variations are unclear.

Kazhak et al. [8] have found high expression of hsRPB7 mRNA in heart, liver, skeletal muscle, kidney and low expression in brain, placenta and lung. Although we find high hsRPB7 expression in heart and kidney, we also find high expression in brain and retina. Our blots also show moderate expression in lung, skeletal muscle and placenta, but very low expression in liver and stomach (Fig. 5). These discrepancies may be due to hsRPB7 mRNA instability in some tissues or to individual or circadian variations. These types of variables are difficult to control in human samples. Another possibility is that hsRPB7 modulates the Pol II complex to preferentially transcribe certain types of genes more actively in certain tissues. This might explain the low levels found in adult liver even though liver is known to actively transcribe a wide variety of genes. However, an argument could be made that hsRPB7 mRNA levels are just a reflection of the level of transcription. For example, skeletal muscle has a hsRPB7 expression which is approximately half that of heart muscle and only 20–25% of the level of expression found in brain and retina. These other tissues are likely to have higher metabolic and transcription rates than most skeletal muscle. There are also developmental differences in expression (Fig. 6). Interestingly, fetal liver has a moderate levels of expression but it is still only about 20% of that found in fetal brain.

The 5′ flanking region or putative promoter of hsRPB7 (Fig. 9) lack a TATA box but contains several Sp1 sites and a AP1 site. This is consistent with a constitutive or 'housekeeping' promoter and may explain the wide-ranging expression. Although we are not providing any direct evidence that the sequence in Fig. 9 is an active promoter, our 5′ RACE results and genomic sequence analysis strongly suggests that this is the hsRPB7 promoter. Other possible transcription factor binding sites are present within this region but until functional analyses are performed, it is difficult to speculate as to their importance in transcribing this gene.

Khazak et al. [8], have shown that hsRPB7 has 63% amino acid similarity to the yeast RPB7 and can functionally complement yeast RPB7 deletion strains. These results suggest that hsRPB7 would be highly conserved in eukaryotic species. We have performed a multispecies genomic blot (Fig. 10) to see the extent of conservation among different eukaryotic species. Our results suggest that RPB7 is highly conserved at the nucleic acid level among eukaryotic species, especially mammals. These result also indicate that there is only one RPB7 gene in most species.

The RNA polymerase II complex is involved in a variety of complex interactions as it performs its main function, transcribing expressed genes. Another well known function of this complex is DNA repair [17] and another suspected function is the maintenance of the chromatin structure [18]. In the past few years it has become clear that the RNA Pol II complex and its involvement in transcription coupled DNA repair plays an important role in human disease [17,19]. Two rare but highly heterogeneous human genetic disorders have been shown to be caused by defects in genes whose products are involved in transcription-coupled DNA repair, xeroderma pigmentosum [20] and Cockayne syndrome [21]. However, defects in the transcription-coupled DNA repair mechanism are suspected in a broader range of diseases such as those of a neurodegenerative nature [22] as well as cancer [23]. The information provided by this study will facilitate further work to determine if hsRPB7 is involved in some of these above-mentioned diseases.

## References

[1] M. Sawadogo, A. Sentenac, Annu. Rev. Biochem. 59 (1990) 711–754.
[2] S. Buratowski, Cell 77 (1994) 1–3.
[3] J. Acker, M.-G. Mattei, M. Wintzerith, N. Roeckel, D. Depetris, M. Vigneron, C. Kedinger, Genomics 20 (1994) 496–499.
[4] K. Mita, H. Tsuji, M. Morimyo, E. Takahashi, M. Nenoi, S. Ichimura, M. Yamauchi, E. Hongo, A. Hayashi, Gene 159 (1995) 285–286.
[5] J. Acher, M. Wintzerith, M. Vigneron, C. Kedinger, J. Seq. Map. 4 (1994) 328–331.
[6] T.J. Schoen, K. Mazuruk, G.J. Chader, I.R. Rodriguez, Biochem. Biophys. Res. Commun. 213 (1995) 181–188.
[7] S.F. Altschul, W. Gish, W. Miller, E. Myers, D. Lipman, J. Mol. Biol. 215 (1990) 403–410.
[8] V. Khazak, P.P. Sadhale, N.A. Woychik, R. Brent, E.A. Golemis, Mol. Biol. Cell 6 (1995) 759–775.
[9] I.R. Rodriguez, K. Mazuruk, T.J. Schoen, G.J. Chader, J. Biol. Chem. 269 (1994) 31969–31977.
[10] D.B. Showalter, S.S. Sommer, Anal. Biochem. 177 (1989) 90–94.

[11] D.S. Gerhard, E. Lawrence, J. Wu, H. Chua, N. Ma, S. Bland, C. Jones, Genomics 13 (1992) 1133–1142.

[12] M. Lizuka, C. Jones, K. Hayashi, T. Sekiya, Genomics 19 (1994) 581–584.

[13] P. Couillin, E. Le Guern, A. Vignal, C. Fizames, N. Ravise, D. Delportes, I. Reguigne, M.F. Rosier, C. Junien, V. Van Heyningen, Weissenbach, Genomics 21 (1994) 379–387.

[14] H. Stohr, B.H. Weber, Hum. Genet. 95 (1995) 219–222.

[15] B.H. Weber, G. Vogt, R.C. Pruett, H. Stohr, H. Felbor, Nat. Genet. 8 (1994) 352–356.

[16] A. Meindl, K. Dry, K. Herrman, F. Manson, A. Ciccodicola, A. Edgar, M.R. Carvalho, H. Achatz, H. Hellebrand, A. Lennon, C. Migliaccio, K. Porter, E. Zrenner, A. Bird, M. Jay, B. Lorenz, B. Wittwer, M. D'Urso, T. Meitinger, A. Wright, Nat. Genet. 13 (1996) 35–42.

[17] P.C. Hanawalt, Science 266 (1994) 1957–1958.

[18] P.R. Cook, BioEssays 16 (1994) 425–430.

[19] M. Digweed, Toxicol. Lett. 67 (1993) 259–281.

[20] K.H. Kraemer, D.D. Levy, C.N. Parris, E.M. Gozukara, S. Moriwaki, S. Adelberg, M.M. Seidman, J. Invest. Dermatol. 103 (supplement) (1994) 96S–101S.

[21] K.A. Henning, L. Li, N. Iyer, L.D. McDaniel, M.S. Reagan, R. Legerski, R.A. Schultz, M. Stefanini, A.R. Lehmann, L.V. Mayne, E.C. Friedberg, Cell 82 (1995) 555–564.

[22] P. Mazzarello, M. Poloni, S. Spadari, F. Focher, J. Neurol. Sci. 112 (1992) 4–14.

[23] A.M.R. Taylor, C.M. McConville, P.J. Byrd, Br. Med. Bullet. 50 (1994) 708–717.