Localization and Characterization of the Human ADP-Ribosylation Factor 5 (ARF5) Gene

Rachel E. McGuire,^{*,†} Stephen P. Daiger,^{*,‡} and Eric D. Green^{†,1}

^{*}Human Genetics Center and [†]Department of Ophthalmology, University of Texas Health Science Center, Houston, Texas 77225; and [‡]Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892

Received October 21, 1996; accepted February 12, 1997

ADP-ribosylation factor 5 (ARF5) is a member of the ARF gene family. The ARF proteins stimulate the *in vitro* ADP-ribosyltransferase activity of cholera toxin and appear to play a role in vesicular trafficking *in vivo*. We have mapped ARF5, one of the six known mammalian ARF genes, to a well-defined yeast artificial chromosome contig on human chromosome 7q31.3. In addition, we have isolated and sequenced an ~3.2-kb genomic segment that contains the entire ARF5 coding region, revealing the complete intron-exon structure of the gene. With six coding exons and five introns, the genomic structure of ARF5 is unique among the mammalian ARF genes and provides insight about the evolutionary history of this ancient gene family. © 1997 Academic Press

The ADP-ribosylation factor (ARF) family is a subgroup of the ras superfamily, composed of small (~20-kDa) guanine-nucleotide binding proteins (3, 6). In the presence of GTP, ARF proteins stimulate the *in vitro* ADP-ribosyltransferase activity of cholera toxin, leading to ADP-ribosylation of the α -subunit of G-proteins and the activation of adenylyl cyclase (12). While the *in vivo* function(s) of ARF proteins is not well defined, their physical association with the Golgi apparatus suggests a role in vesicular transport of proteins (13, 17).

To date, six mammalian ARF genes have been identified (2, 9, 11, 15, 20), all of which contain identical sequence motifs for guanine nucleotide-binding and GTP hydrolysis. All of the ARF genes studied appear to be widely expressed in human tissues. Based on their sequence, size, and phylogeny, the mammalian ARF proteins have been divided into three classes: Class I (ARF1, ARF2, and ARF3), Class II (ARF4 and ARF5), and Class III (ARF6) (20). The ARF genes are

highly conserved, being found in species as divergent as human, yeast, and *Giardia*. From the predicted amino acid sequence, human ARF1 is 100% identical to bovine ARF1 and human ARF6 is 99% identical to CPS1, a chicken pseudogene (1, 2). Human transcripts have been isolated for all but ARF2 (2, 9, 18–20); however, the latter has been found in bovine, rat, and mouse (19).

Recently, the first human ARF genes were localized in the genome: ARF1 and ARF3 were mapped to chromosomes 1q42 and 12q13, respectively (5). In addition, an ARF-like gene, ARF4L, was mapped to chromosome 17q12–q21 (16). To date, none of the ARF genes has been associated with human disease. We recently constructed a yeast artificial chromosome (YAC)-based sequence-tagged site (STS)-content map of the critical region for RP10, one of eight genes that cause autosomal dominant retinitis pigmentosa (8). This region maps to human chromosome 7q31.3. A number of expressed-sequence tags (ESTs) were localized within this contig. Subsequently, we analyzed the ESTs more rigorously to evaluate their possible candidacy as the RP10 gene and discovered that one of these (EST20241, GenBank T27924, GDB:591767 and GDB:589476) corresponded to the human ARF5 gene. Here, we report the genomic sequence, intron-exon structure, and expression pattern of the human ARF5 gene.

A cosmid clone (161C9) containing the ARF5 gene was isolated from a chromosome 7-specific cosmid library prepared from flow-sorted DNA (Lawrence Livermore National Laboratory, LL07NC01) by hybridization and subsequent PCR confirmation [using the PCR assay corresponding to the ARF5-specific STS (see GDB:589476)]. Based on the established cDNA sequence of human ARF5 (GenBank Accession No. M57567), numerous primers were designed scattered along the entire length of the coding sequence. These primers were then used in a pair-wise fashion for performing PCR with the cosmid clone as template. Fortunately, a pair of primers at the far 5' (ATGGGCCTC-ACCGTGTCCGCG) and 3' (TCCAACAACAGAGAG-AACATGC) ends, respectively, of the ARF5 cDNA sequence was found to amplify an ~3.2-kb product. This

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. U73002.

¹To whom correspondence should be addressed at National Human Genome Research Institute, National Institutes of Health, 49 Convent Drive, MSC4431, Building 49, Room 2A08, Bethesda, MD 20892. Telephone: (301) 402-0201. Fax: (301) 402-4735. E-mail: egreen@nhgri.nih.gov.

[illegible]

FIG. 1. Genomic sequence of the human ARF5 gene. The DNA sequence of the coding regions (uppercase letters) as well as the five introns and flanking regions (lowercase letters) of the human ARF5 gene are shown. The predicted 160-amino-acid sequence is indicated in boldface, the four guanine nucleotide-binding consensus sequences are underlined, and the polyadenylation signal (AATAAA) is double-underlined. This sequence was deposited in GenBank (Accession No. U73002).

fragment was cloned into pCRII (Invitrogen) and fully sequenced using the above cDNA-specific primers (as well as additional primers designed from newly obtained sequence) and dye-labeled terminators (Perkin-Elmer/Applied Biosystems). The resulting 3240-bp genomic sequence, which contains the entire ARF5 coding sequence, is provided in Fig. 1 (see GenBank Accession No. U73002). Direct comparison of the genomic sequence and cDNA sequence allowed the intron-exon boundaries of the human ARF5 gene to be defined, revealing the presence of six exons and five introns.

The genomic structure of human ARF5 is quite dif-

ferent than the other ARF genes characterized to date. For example, the Class I ARF genes contain four exons, five introns, and identical intron-exon boundaries (14). Thus, the intron-exon organization of human ARF5 is notably different than the Class I genes. From an evolutionary perspective, these results are potentially of interest. In 1994, Lee *et al.* (7) constructed an evolutionary tree of the ARF family using the cDNA-deduced amino acid sequences of the human, bovine, *Drosophila*, yeast, and *Giardia* ARF genes (7). These results suggested that the three classes of ARF genes split before the divergence of insects. In addition, the *Dro-*

FIG. 2. Structural comparison of ARF genes. Schematic representations of the indicated human (h), *Drosophila* (d), and bovine (b) ARF gene structures are depicted. Coding exons and introns are represented by thick and thin lines, respectively. The exons (but not the introns) are drawn to scale. The numbers above each exon reflect the corresponding base positions in the cDNA sequence.

sophila ARFI shares >93% amino acid identity with the human and bovine Class I ARF proteins (10). *Drosophila* ARFII is 84% identical at an amino acid level to human ARF4 and ARF5 and is thus phylogenetically grouped into Class II (7). Figure 2 shows the genomic structure and alignment of human ARF5, *Drosophila* ARFII, human and bovine Class I, and the *Drosophila* ARFI genes. Human ARF5 and *Drosophila* ARFII have similar intron–exon boundaries, though the *Drosophila* ARFII exon 2 is split into exons 2, 3, and 4 in human ARF5. These findings reveal fundamentally different structural organizations for Class I and II genes and indicate that these structural differences are ancient, as they are also found in *Drosophila*.

Northern hybridization analysis of ARF5 revealed

high-level expression of the ~1-kb message in human skeletal muscle and pancreas, lower-level expression in heart, brain, placenta, spleen, thymus, prostate, testis, small intestine, colon, and leukocytes, and little to no expression in lung, liver, kidney, and ovary (data not shown). However, PCR analysis detected ARF5 sequences in a number of cDNA libraries, specifically, those constructed from retina, placenta, fetal brain, HeLa cell, and thymus. In evaluating the candidacy of ARF5 as the cause of RP10, individual exons were PCR amplified and sequenced in two affected and three unaffected individuals from a previously reported RP10 family (8). The primers used to amplify each exon individually are provided in Table 1. No mutations were encountered in RP10-affected individuals.

TABLE 1
Primers for PCR Amplification of Individual Human ARF5 Exons

Exon	Primer name	Primer sequence ^a	PCR product size (bp) ^b
1	ARF5-1F	(TGTAACACGACGGCCAGT) CGCGTCGGTGCCCGCG	252
	ARF5-1R	(CAGGAAACAGCTATGACC) ACCCGCCGAGAGCAGAG	
2	ARF5-2F	(TGTAACACGACGGCCAGT) CCTTTCCTGGTCTCTAGGG	179
	ARF5-2R	(CAGGAAACAGCTATGACC) TCACCTATGGTTGGGATGG	
3	ARF5-3F	(TGTAACACGACGGCCAGT) GTTTGTGTCCTGCTGAAATC	231
	ARF5-3R	(CAGGAAACAGCTATGACC) GTTAGAAAGTCCCCTCCCAAC	
4	ARF5-4F	(TGTAACACGACGGCCAGT) TATGTCCAGGCAGAAGTGA	173
	ARF5-4R	(CAGGAAACAGCTATGACC) ACTCTGTCCCCAAGCTGAGG	
5	ARF5-5F	(TGTAACACGACGGCCAGT) AGTTCCTCAACACAGTAGAG	224
	ARF5-5R	(CAGGAAACAGCTATGACC) AGGCAGGATTCAGCACA	
6	ARF5-6F	(TGTAACACGACGGCCAGT) GGCATTCTTTTGTTCCTG	154
	ARF5-6R	(CAGGAAACAGCTATGACC) CCTGCCCTGGCTGGTTA	

^a All primers included one of two common 18mer sequences at the 5' end (indicated within parentheses) that subsequently provided an annealing site for fluorescently labeled universal sequencing primers (4).

^b Each PCR product size reflects the amplification of ARF5 sequences only and does not include the extra 36 bp contributed by the extra 18mer tail on each primer.

In summary, we have established the chromosomal location, genomic sequence, and intron-exon structure of the human ARF5 gene. The entire coding region is small, being contained within ~3.2 kb of genomic DNA, and is divided among six exons. Furthermore, in conjunction with previous studies (8), the gene is now localized on a well-established YAC contig from cytogenetic band 7q31.3 and maps in close physical proximity to Génethon genetic marker AFM206xc1 (D7S635,GDB:199240).

ACKNOWLEDGMENTS

We thank Dr. Joel Moss (National Heart, Lung, and Blood Institute, National Institutes of Health) for helpful advice and technical assistance in characterizing the human ARF5 sequence; Dr. Gerry Bouffard for computational assistance; and Drs. Darrell Ellsworth, Gerry Bouffard, Jeff Touchman, Joel Moss, Martha Vaughan, and Walter Patton for helpful comments about the manuscript. This work was supported in part by grants from the Foundation Fighting Blindness and the George Gund Foundation (to S.P.D.), by NIH-NEI Grant EY07142 (to S.P.D.), and by Institutional National Research Service Award EY07024 (to R.E.M.).

REFERENCES

- Alsip, G. R., and Konkel, D. A. (1986). A processed chicken pseudogene (CPS1) related to the ras oncogene superfamily. *Nucleic Acids Res.* **14**: 2123–2138.
- Bobak, D. A., Nightingale, M. S., Murtagh, J. J., Price, S. R., Moss, J., and Vaughan, M. (1989). Molecular cloning, characterization, and expression of human ADP-ribosylation factors: Two guanine nucleotide-dependent activators of cholera toxin. *Proc. Natl. Acad. Sci. USA* **86**: 6101–6105.
- Bobak, D. A., Tsai, S.-C., Moss, J., and Vaughan, M. (1990). In "ADP-Ribosylating Toxins and G Proteins: Insights into Signal Transduction" (J. Moss and M. Vaughan, Eds.), pp. 439–456, Am. Soc. Microbiol., Washington, DC.
- Chadwick, R. B., Conrad, M. P., McGinnis, M. D., Johnston-Dow, L., Spurgeon, S. L., and Kronick, M. N. (1996). Heterozygote and mutation detection by direct automated fluorescent DNA sequencing using a mutant Taq DNA polymerase. *Bio-Techniques* **20**: 676–683.
- Hirai, M., Kusuda, J., and Hashimoto, K. (1996). Assignment of human ADP ribosylation factor (ARF) genes ARF1 and ARF3 to chromosomes 1q42 and 12q13, respectively. *Genomics* **34**: 263–265.
- Kahn, R. A., and Gilman, A. G. (1984). Purification of a protein cofactor required for ADP-ribosylation of the stimulatory regulatory component of adenylate cyclase by cholera toxin. *J. Biol. Chem.* **259**: 6228–6234.
- Lee, F.-J. S., Stevens, L. A., Hall, L. M., Murtagh, J. J., Kao, Y. L., Moss, J., and Vaughan, M. (1994). Characterization of class I and class II ADP-ribosylation factor genes and proteins in *Drosophila melanogaster*. *J. Biol. Chem.* **269**: 21555–21560.
- McGuire, R. E., Jordan, S. A., Braden, V. V., Bouffard, G. G., Humphries, P., Green, E. D., and Daiger, S. P. (1996). Mapping the RP10 locus for autosomal dominant retinitis pigmentosa on 7q: Refined genetic positioning and localization within a well-defined YAC contig. *Genome Res.* **6**: 255–266.
- Monaco, L., Murtagh, J. J., Newman, K. B., Tsai, S.-C., Moss, J., and Vaughan, M. (1990). Selective amplification of an mRNA and related pseudogene for a human ADP-ribosylation factor, a guanine nucleotide-dependent protein activator of cholera toxin. *Proc. Natl. Acad. Sci. USA* **87**: 2206–2210.
- Murtagh, J. J., Lee, F.-J. S., Deak, P., Hall, L. M., Monaco, L., Lee, C.-M., Stevens, L. A., Moss, J., and Vaughan, M. (1993). Molecular characterization of a conserved, guanine nucleotide-dependent ADP-ribosylation factor in *Drosophila melanogaster*. *Biochemistry* **32**: 6011–6018.
- Price, S. R., Nightingale, M., Tsai, S.-C., Williamson, K. C., Adamik, R., Chen, H.-C., Moss, J., and Vaughan, M. (1988). Guanine nucleotide-binding proteins that enhance cholera ADP-ribosyltransferase activity: Nucleotide and deduced amino acid sequence of an ADP-ribosylation factor cDNA. *Proc. Natl. Acad. Sci. USA* **85**: 5488–5491.
- Schleifer, L. S., Kahn, R. A., Hanski, E., Northrup, J. K., Sternweis, P. C., and Gilman, A. G. (1982). Requirements for cholera toxin-dependent ADP-ribosylation of the purified regulatory component of adenylate cyclase. *J. Biol. Chem.* **257**: 20–23.
- Serventi, I. M., Moss, J., and Vaughan, M. (1992). In "Current Topics in Microbiology and Immunology" (K. Aktories, Ed.), Vol. 175, pp. 43–67, Springer-Verlag, Heidelberg, Germany.
- Serventi, I. M., Cavanaugh, E., Moss, J., and Vaughan, M. (1993). Characterization of the gene for ADP-ribosylation factor (ARF) 2, a developmentally regulated, selectively expressed member of the ARF family of ~20-kDa guanine nucleotide-binding proteins. *J. Biol. Chem.* **268**: 4863–4872.
- Sewell, J. L., and Kahn, R. A. (1988). Sequences of the bovine and yeast ADP-ribosylation factor and comparison to other GTP-binding proteins. *Proc. Natl. Acad. Sci. USA* **85**: 4620–4624.
- Smith, S. A., Holik, P. R., Stevens, J., Melis, R., White, R., and Albertsen, H. (1995). Isolation and mapping of a gene encoding a novel human ADP-ribosylation factor on chromosome 17q12–q21. *Genomics* **28**: 113–115.
- Stearns, T., Willingham, M. C., Botstein, D., and Kahn, R. A. (1990). ADP-ribosylation factor is functionally and physically associated with the Golgi complex. *Proc. Natl. Acad. Sci. USA* **87**: 1238–1242.
- Tsai, S.-C., Adamik, R., Tsuchiya, M., Chang, P. P., Moss, J., and Vaughan, M. (1991). Differential expression during development of ADP-ribosylation factors, 20-kDa guanine nucleotide-binding protein activators of cholera toxin. *J. Biol. Chem.* **266**: 8213–8219.
- Tsuchiya, M., Price, S. R., Nightingale, M. S., Moss, J., and Vaughan, M. (1989). Tissue and species distribution of mRNA encoding two ADP-ribosylation factors, 20-kDa guanine nucleotide binding proteins. *Biochemistry* **28**: 9668–9673.
- Tsuchiya, M., Price, S. R., Tsai, S.-C., Moss, J., and Vaughan, M. (1991). Molecular identification of ADP-ribosylation factor mRNAs and their expression in mammalian cells. *J. Biol. Chem.* **266**: 2772–2777.