A Concern Regarding the Current Confusion with the Human Homolog of Mouse Np95, ICBP90/UHRF1

Authors: Unoki, Motoko, Bronner, Christian, and Mousli, Marc

Source: Radiation Research, 169(2) : 240-244

Published By: Radiation Research Society

URL: https://doi.org/10.1667/RR1209.1

RADIATION RESEARCH **169**, 240–244 (2008)
0033-7587/08 $15.00

COMMENTARY

A Concern Regarding the Current Confusion with the Human Homolog of Mouse Np95, ICBP90/UHRF1

Motoko Unoki,[a,1] Christian Bronner[b] and Marc Mousli[c]

[a] *Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland 20892-4258;* [b] *CNRS UMR 7175, Département de Pharmacologie et Pharmacochimie des interactions moléculaires et cellulaires, Faculté de Pharmacie, 67401 Illkirch Cedex, France; and* [c] *Institut de Parasitologie et de Pathologie Tropicale de Strasbourg, UPRES E.A. 3950 Interactions Cellulaires et Moléculaires Hôte-Parasite, Faculté de Médecine, Université Louis Pasteur, 67000 Strasbourg, France*

Unoki, M., Bronner, C. and Mousli, M. A Concern Regarding the Current Confusion with the Human Homolog of Mouse Np95, ICBP90/UHRF1. *Radiat. Res.* 169, 240–244 (2008).

ICBP90/UHRF1, which is overexpressed in cancer cells and is down-regulated by p53, possesses a methylated CpG binding affinity and binds to the methylated promoters of tumor suppressor genes in cancer cells with HDAC1 and DNMT1, suggesting suppression of these genes and maintenance of methylation status which leads to carcinogenesis. Recently, it was reported that the human homolog of *Np95* is different from *ICBP90* but not from *UHRF1*. Because *UHRF1* is the gene symbol of *ICBP90*, the claim is a little confusing; that is, *UHRF1* and *ICBP90* are identical. Because the previously published genomic structure of the *ICBP90* gene needed to be revised and the registered *ICBP90* sequence (AF129507) contains two rare polymorphisms or sequence errors, we think that confusion could occur. Here we show the revised *ICBP90* gene structure and 366 polymorphisms in this gene. Our conclusion is that the human homolog of Np95 is ICBP90, whose gene symbol is *UHRF1*.

INTRODUCTION

ICBP90 (gene symbol *UHRF1*), which was first identified as a *topoisomerase II alpha* regulator, is a very important protein that is associated with cell cycle progression and is overexpressed in proliferating cells and cancer cells (*1–3*). ICBP90 has a binding affinity to methylated CpGs, makes a complex with HDAC1, and binds to the promoters of methylated tumor suppressor genes, suggesting that ICBP90 promotes cancer cell proliferation through suppression of these tumor suppressor genes (*2*). ICBP90 is also down-regulated by p53 (*4*), possesses ubiquitin E3 ligase activity (*5*), and plays a role in maintaining DNA methylation with DNMT1, keeping the genes silent (*6, 7*). ICBP90 has three family proteins, NIRF/UHRF2, ICBP55/UHRF3 and ICBP87/UHRF4, that share 52.6, 59.8 and 78.1% identity with ICBP90, respectively (*3*).

Recently an article titled ''Isolation and Characterization of a Novel Human Radiosusceptibility Gene, *NP95*'' was published in *Radiation Research* (*8*). The authors claimed that they identified a new human homolog of mouse Np95 and the ''new'' gene was identical to *UHRF1* but not to *ICBP90.* However, because the gene symbol of ICBP90 is *UHRF1* (i.e., they are the same) and the *ICBP90* sequence has been registered for a long time (AF129507, submitted by Hopfner *et al.* on Feb. 19, 1999), we are afraid that the article will create major confusion in the future. We would like to discuss the possibilities that could have led to the confusion.

MATERIALS AND METHODS

Databases

We used the databases below to determine the genomic DNA structure of *ICBP90* and polymorphisms in the gene: BLAST (http://www.ncbi.nlm.nih.gov/BLAST/), Unigene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene), Ensemble (http://www.ensembl.org/index.html), GenBank (http://www.ncbi.nlm.nih.gov/Genbank/), DDBJ (http://www.ddbj.nig.ac.jp/), dbSNP (http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp), and JSNP (http://snp.ims.u-tokyo.ac.jp/).

RESULTS AND DISCUSSION

Concerning the recently published article entitled ''Isolation and Characterization of a Novel Human Radiosusceptibility Gene, NP95'' (*8*), we have a strong concern that the authors did not isolate a new gene. In the discussion of the article, it is described that they could not recognize the sequence of exon F (2.6 kb) of the *ICBP90* gene in the *Homo sapiens* chromosome 19 genomic contig, NT_011255 (GenBank), at 19p13.3, but found this exon in a BAC

[1] Address for correspondence: Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, NIH, 37 Convent Drive, Building 37, Room 3060, Bethesda, Maryland 20892-4258; e-mail: unokim@mail.nih.gov.

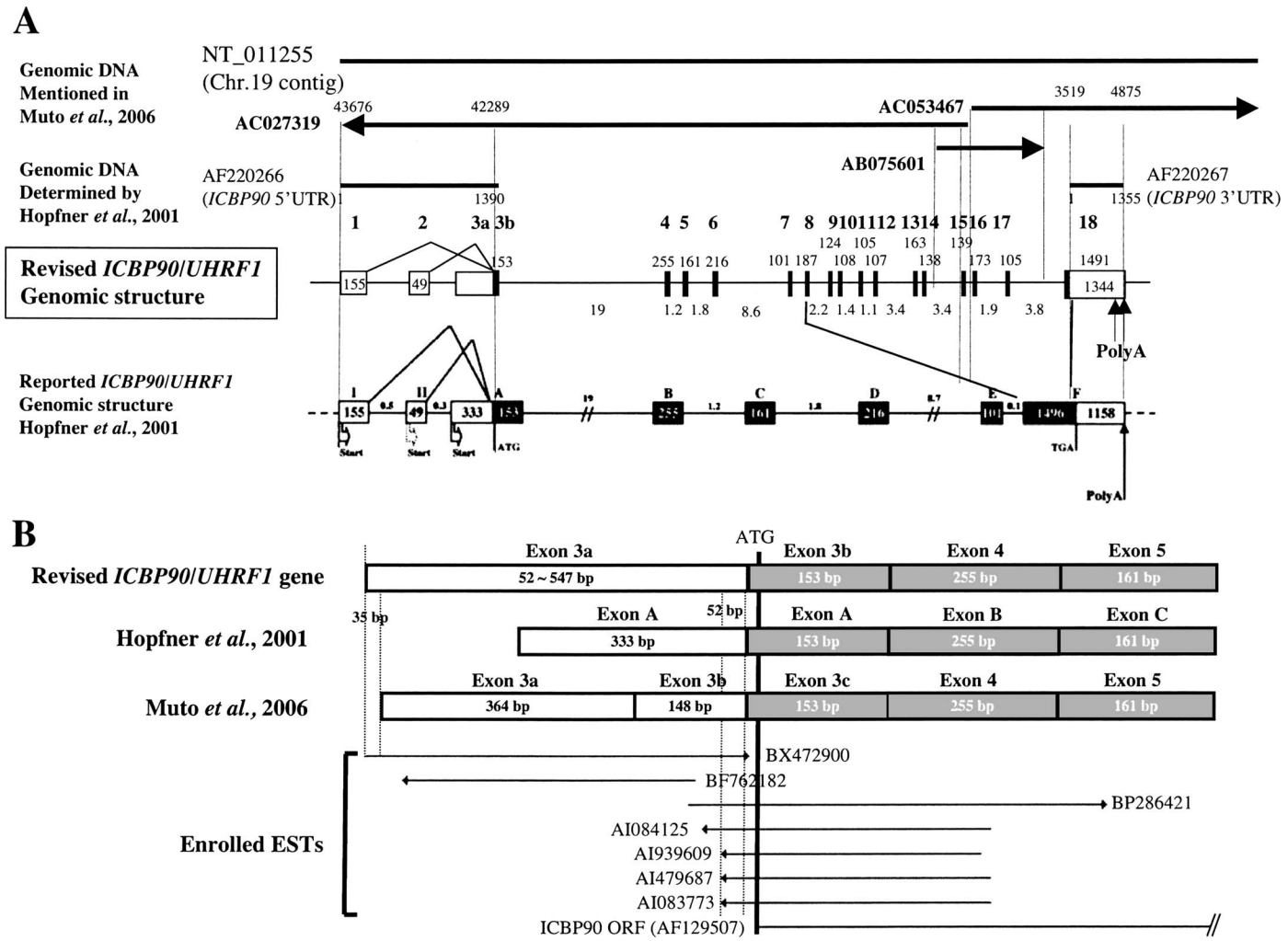

FIG. 1. Panel A: Revised full-length *ICBP90/UHRF1* gene on chromosome 19 contig, NT_01125. AC027319, AC053467 and AB075601 are the three genomic DNA fragments that were mentioned by Muto *et al.* (*8*). Exon F reported by Hopfner *et al.* (*9*) was composed of 11 exons. AF220266 and AF220267 were the *ICBP90* 5′UTR and 3′UTR sequence determined by Hopfner *et al.* (*9*). Panel B: Exon 3 is composed of exons 3a and 3b. The length of exon 3a is 52–547 bp based on the seven registered ESTs. Exon 3b is 153 bp including the start codon (ATG) and is involved in all transcripts.

clone, AC112777, that includes 12p12.2-p12.1. Hopfner *et al.* (*9*) have reported that the *ICBP90* gene spans approximately 35.8 kb on chromosome 19p13.3 and contains six coding exons named A to F. We carefully checked the genomic structure of the *ICBP90* gene and found that exon F is actually composed of 11 exons (Fig. 1). All exons of *ICBP90* are on the contig, NT_011255, which is composed of many genomic fragments, including AC027319, AC053467 and AB07561, which are discussed in the article. Exons 1 to 14 are on AC027319, exons 15 to 17 are on AB075601, and exons 16 to 18 are on AC053467. Therefore, the *ICBP90* gene is localized at the same position where the ''new'' human *NP95* homolog is located. We think that this incorrect definition of exon F (*9*) is one of the major sources of confusion that led Muto *et al.* to think that the *ICBP90* gene is different from the gene that they identified.

It is also stated in their discussion that human *NP95* and *UHRF1* are identical in ORF and that *ICBP90* and human *NP95* differ in ORF by only two amino acids (Lys383 → Asn383, Ala457 → Ser457) and seven nucleotides. A difference of nine nucleotides in a long identical sequence (in this case, *ICBP90*'s ORF is 2382 bp) does not make it a different gene. Usually these differences are polymorphisms or sequence errors. Therefore, we searched SNP databases and found 366 polymorphisms including 308 single nucleotide polymorphisms (SNPs) and 58 deletions on the gene (Table 1, Supplemental Table 1). Among these 366 polymorphisms, 337 were located in introns, six were in 3′UTR, one was in 5′UTR, and 22 were in exons (Table 1). Among the 22 polymorphisms in exons, 15 were synonymous substitutions and seven were nonsynonymous substitutions (Table 2).

We are not sure whether all nine nucleotide differences are polymorphisms, because there are no descriptions of seven of the nine different nucleotides in the article. How-

TABLE 1
Polymorphisms in the *ICBP90/UHRF1* Gene

Location	Deletion	Substitution (SNP)	Synonymous substitutions	Nonsynonymous substitutions	Total polymorphism numbers
Intron	58	279	0	0	337
3′UTR	0	6	0	0	6
5′UTR	0	1	0	0	1
Coding region	0	22	15	7	22
Total	58	308	15	7	366

ever, the two differences (Lys383 → Asn383, Ala457 → Ser457) that are described in the article were not involved in the 366 polymorphisms. We determined the sequence by sequencing genomic DNA from several cancer cell lines and Japanese individuals and determined that codon 383 is AAG (Lys) and codon 457 is GCG (Ala); on the other hand, codon 383 is AAT (Asn) and codon 457 is TCG (Ser) in the registered *ICBP90* sequence (AF129507). Therefore, AF129507 includes at least two rare polymorphisms that are not registered in the database or are sequence errors at codons 383 and 457. The other seven different nucleotides could be polymorphisms. Unoki *et al.* (*2*) used *ICBP90* plasmid vectors that include Lys383 and Ala457, not Asn383 and Ser457, and three different polymorphisms compared with AF129507, because these polymorphisms were major alleles in the database. If the claim of Muto *et al.* (*8*) is right, Unoki *et al.* (*2*) did not use ICBP90 in their experiments. But no one accepts that two genes at exactly the same location in a chromosome with 99.748% amino acid identity (only two amino acids are different among 754 amino acids) are different.

Our other concern is the 5′UTR splice variations that are described in Fig. 2 (*8*). Hopfner *et al.* (*9*) reported three 5′UTR variations that are different from the variations in their article. We checked the EST database using BLAST and found that the three different 5′UTR variations that Hopfner *et al.* (*9*) described were registered as EST fragments, indicating that they are actually expressed in cells (Fig. 2, Supplemental Table 2). However, there are no ESTs that include exon 1 and 2 as described by Muto *et al.* (*8*). Transcripts that include these exons may be expressed at a low level, but we need expression data. At the least, they should add the 5′UTR variation that Hopfner *et al.* (*9*) identified in their article.

The authors also claimed that their newly isolated gene differs in its effect on topoisomerase II activity. However,

TABLE 2
Detailed SNP Information in *ICBP90/UHRF1* Gene Exons

dbSNP rs number		Flanking sequence	Exon location
15 Synonymous SNPs			
1	rs2261986	TGTGGATCCAGGTTCGGACCATGGA[C/T]GGGAGGCAGACCCACACGGTGGACTC	exon 3
2	rs2251520	GAGTCAGACAAGTCCTCCACCCACGG[T/C]GAGGCGGCCGCCGAGACTGACAGCA	exon 4
3	rs2123731	GATGAGGACATGTGGGATGAGACGGA[A/G]TTGGGGCTGTACAAGGTGAGCCTCC	exon 4
4	rs2307205	TCCCGGGACGAGCCCTGCAGCTCCAC[G/A]TCCAGGCCGGCGCTGGAGGAGGACG	exon 5
5	rs2307201	GGCGTGGTCCAGATGAACTCCAGGGA[C/T]GTCCGAGCGCGCGCCCGCACCATCA	exon 6
6	rs2307206	CAGATGAACTCCAGGGACGTCCGAGC[G/A]CGCGCCCGCACCATCATCAAGTGGC	exon 6
7	rs17881281	CTGGAGGTGGGCCAGGTGGTCATGCT[C/T]AACTACAACCCCGACAACCCCAAGG	exon 6
8	rs4807002	GCCCCGTGCCCAGGGAAGAGCGGGCC[G/A]TCCTGCAAGCACTGCAAGGACGACG	exon 8
9	rs17883422	GACCCGCCCCTCAGCAGTGTTCCCAG[C/T]GAGGACGAGTGGTGAGTGCGGCCCT	exon 8
10	rs2250982	GTCCATCGGCCCCACGTGGCTGGCAT[C/A]CACGGCCGGAGCAACGACGGAGCGT	exon 11
11	rs2250981	CATCGGCCCCACGTGGCTGGCATACA[C/T]GGCCGGAGCAACGACGGAGCGTACT	exon 11
12	rs2250978	GTCCTGGCGGGGGGCTATGAGGATGA[C/T]GTGGTGAGTGTGTGTGTGGGAGGGG	exon 11
13	rs2307213	TACGCCCCCGCTGAGGGCAACCGCTA[C/T]GATGGCATCTACAAGGTGAGTGCCC	exon 13
14	rs17878253	AGGGAGGAGGAGGAGCAGCAGGAGGG[G/T]GGCTTCGCGTCCCCCAGGACGGGCA	exon 15
15	rs2247238	GAGGAGGAGCAGCAGGAGGGGGGCTT[C/T]GCGTCCCCCAGGACGGGCAAGGGCA	exon 15
7 Nonsynonymous SNPs			
1	rs17886098	ACCCCAAGGAGCGGGGCTTCTGGTAC[G/C]ACGCGGAGATCTCCAGGAAGCGCGA	exon 6
2	rs2292148	GCCGTCCTGCAAGCACTGCAAGGAC[A/G]ACGTGAACAGACTCTGCCGGGTCTGC	exon 8
3	rs17885791	TGGTACTGGCGGGAGAGCGGCTGAGA[G/A]AGAGCAAGAAGAAGGCGAAGATGGC	exon 9
4	rs2307211	CAAGTACGCCCCCGCTGAGGGCAACC[G/T]CTATGATGGCATCTACAAGGTGAGT	exon 13
5	rs17883331	AGGAGGAGCAGCAGGAGGGGGGCTTC[G/A]CGTCCCCCAGGACGGGCAAGGGCAA	exon 15
6	rs17884843	GGAGGGGGGCTTCGCGTCCCCCAGGA[C/T]GGGCAAGGGCAAGTGGAAGCGGAAG	exon 15
7	rs17883563	ACACCCGCTTCCCTCTAGTTCCAGTT[G/T]TTCCTGAGTAAAGTGGAGGAGACGT	exon 17

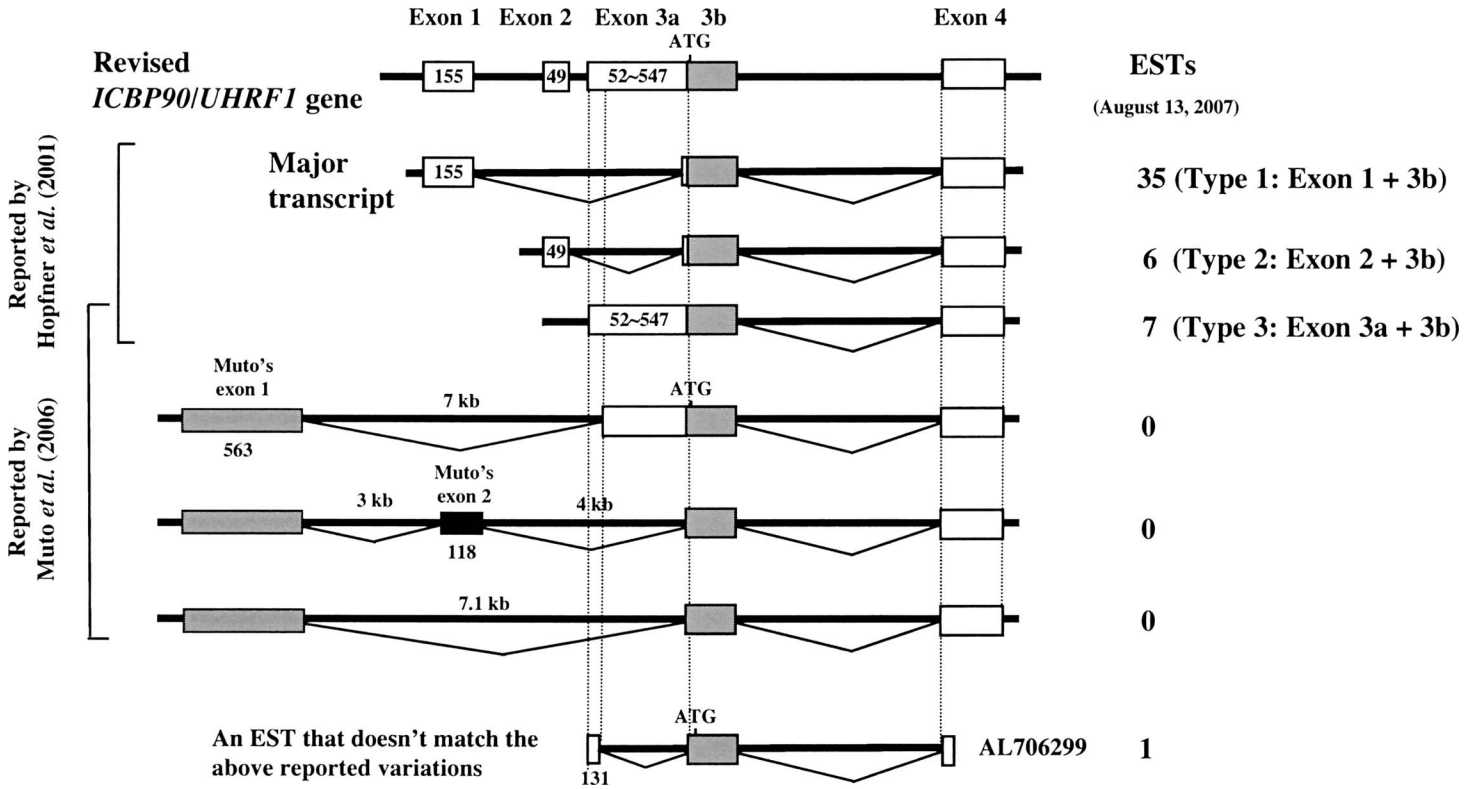

FIG. 2. 5′UTR variations of *ICBP90/UHRF1* transcripts. Major transcript (35 ESTs have been enrolled) has exon 1 and exon 3b. A transcript that has exon 2 and exon 3b (six ESTs have been registered) and another transcript that has exon 3a and exon 3b (seven ESTs have been registered) are secondarily abundant transcripts based on an EST database search. There are no ESTs corresponding to the three transcripts that are described by Muto *et al.* (*8*). AL706299 has a 5′ part of exon 3a and exon 3b and does not match any reported transcripts.

TABLE 2
Extended

Substitution type	Amino acid		Allele frequency		Amino acid position
	Major	Minor	Major	Minor	
synonymous	C: Asp	T: Asp	C: 0.7	T: 0.3	9
synonymous	T: Gly	C: Gly	T: 0.674	C: 0.326	111
synonymous	A: Glu	G: Glu	A: 0.688	G: 0.312	131
synonymous	G: Thr	A: Thr	G: 0.989	A: 0.011	173
synonymous	C: Asp	T: Asp	C: 0.994	T: 0.006	203
synonymous	G: Ala	A: Ala	G: 0.982	A: 0.018	206
synonymous	C: Leu	T: Leu	C: 0.994	T: 0.006	225
synonymous	G: Pro	A: Pro	G: 0.993	A: 0.007	300
synonymous	C: Ser	T: Ser	C: 0.980	T: 0.020	354
synonymous	C: He	A: He	C: 0.591	A: 0.409	449
synonymous	C: His	T: His	C: 0.522	T: 0.478	450
synonymous	C: Asp	T: Asp	C: 0.925	T: 0.075	469
synonymous	C: Tyr	T: Tyr	C: 0.639	T: 0.361	555
synonymous	G: Gly	T: Gly	G: 0.977	T: 0.023	635
synonymous	C: Phe	T: Phe	C: 0.771	T: 0.229	637
nonsynonymous	G: Asp	C: His	G: 0.994	C: 0.006	240
nonsynonymous	G: Asp	A: Asn	G: 0.988	A: 0.012	308
nonsynonymous	G: Glu	A: Lys	G: 0.983	A: 0.017	379
nonsynonymous	G: Arg	T: Leu	G: 0.987	T: 0.013	554
nonsynonymous	G: Ala	A: Thr	G: 0.835	A: 0.165	638
nonsynonymous	C: Thr	T: Met	C: 0.994	T: 0.006	642
nonsynonymous	G: Leu	T: Phe	G: 0.994	T: 0.006	713

the authors did not compare the activity of their ''new'' gene with ICBP90 in the same experiment. The difference between two proteins should be compared in the same experiment. The effect of ICBP90 on topoisomerase II α activation might not be as strong. In fact, Unoki *et al.* (*2*) got a similar result. They showed that ICBP90 has a stronger affinity for methylated CpG than for the CCAAT repeat at the topoisomerase promoter. However, as is well known, to get the maximum protein activity, special conditions are required (salt concentration, binding partner, etc.), and the condition that Unoki *et al.* (*2*) used might not be suitable to get the strongest activation of topoisomerase by ICBP90. Although the results of Muto *et al.* (*8*) are interesting, it is too early to conclude anything.

In conclusion, the genomic structure of *ICBP90* that was defined previously was revised (Fig. 1), and the registered *ICBP90* sequence (AF129507) includes at least two rare polymorphisms or sequence errors. All exons of the *ICBP90* gene are on 19p13.3. *ICBP90* (gene symbol *UHRF1*) is a human homolog of mouse *Np95*, as we have described before (*3*). Actually, because ICBP90 and mouse Np95 share 73.3–73.7% amino acid identity, the percentage is a little low compared to an ordinary homolog between human and mouse. However, because the human genome project was finished in 2003, we do not think that there is another human *NP95* homolog on the human genome. If there is another one, it should share a higher percentage of amino acid identity with mouse Np95 than ICBP90. In addition, functional differences should be compared in the same experiment. Since the gene that the authors isolated and ICBP90 are identical, the authors' functional analysis data seem to suggest that ICBP90's function is the same as Unoki *et al.* reported (*2*) where they used the same sequence that the authors identified. We hope that this commentary can prevent future confusion.

SUPPLEMENTARY INFORMATION

Supplementary Table 1. Detailed information for all SNPs on the *ICBP90/UHRF1* gene. http://dx.doi.org/10.1667/RR1209.1.s1

Supplementary Table 2. ESTs that include *ICBP90/UHRF1* 5′UTR variation (available August 13, 2007). http://dx.doi.org/10/1667/RR1209.1.s2

ACKNOWLEDGMENT

We thank Dr. Tom Holroyd for editorial help.

Received: August 30, 2007; accepted: November 8, 2007

REFERENCES

1. R. Hopfner, M. Mousli, J. M. Jeltsch, A. Voulgaris, Y. Lutz, C. Marin, J. P. Bellocq, P. Oudet and C. Bronner, ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIalpha expression. *Cancer Res.* **60**, 121–128 (2000).
2. M. Unoki, T. Nishidate and Y. Nakamura, ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene* **23**, 7601–7610 (2004).
3. C. Bronner, M. Achour, Y. Arima, T. Chataigneau, H. Saya and V. B. Schini-Kerth, The UHRF family: Oncogenes that are drugable targets for cancer therapy in the near future? *Pharmacol. Ther.* **115**, 419–434 (2007).
4. Y. Arima, T. Hirota, C. Bronner, M. Mousli, T. Fujiwara, S. Niwa, H. Ishikawa and H. Saya, Down-regulation of nuclear protein ICBP90 by p53/p21Cip1/WAF1-dependent DNA-damage checkpoint signals contributes to cell cycle arrest at G_1/S transition. *Genes Cells* **9**, 131–142 (2004).
5. Y. Jenkins, V. Markovtsov, W. Lang, P. Sharma, D. Pearsall, J. Warner, C. Franci, B. Huang, J. Huang and Y. Hitoshi, Critical role of the ubiquitin ligase activity of UHRF1, a nuclear RING finger protein, in tumor cell growth. *Mol. Biol. Cell* **16**, 5621–5629 (2005).
6. M. Bostick, J. K. Kim, P. O. Esteve, A. Clark, S. Pradhan and S. E. Jacobsen, UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
7. M. Achour, X. Jacq, P. Rondé, M. Alhosin, C. Charlot, T. Chataigneau, M. Jeanblanc, M. Macaluso, A. Giordano, A. D. Hughes and C. Bronner, The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of *VEGF* gene expression. *Oncogene*, in press. [doi: 10.1038/sj.onc.1210855]
8. M. Muto, A. Fujimori, M. Nenoi, K. Daino, Y. Matsuda, A. Kuroiwa, E. Kubo, Y. Kanari, M. Utsuno and K. Tatsumi, Isolation and characterization of a novel human radiosusceptibility gene, NP95. *Radiat. Res.* **166**, 723–733 (2006).
9. R. Hopfner, M. Mousli, J. M. Garnier, R. Redon, S. du Manoir, B. Chatton, N. Ghyselinck, P. Oudet and C. Bronner, Genomic structure and chromosomal mapping of the gene coding for ICBP90, a protein involved in the regulation of the topoisomerase II alpha gene expression. *Gene* **266**, 15–23 (2001).