Brief Genetics Report

Structure and Organization of the Human Uncoupling Protein 2 Gene and Identification of a Common Biallelic Variant in Caucasian and African-American Subjects

George Argyropoulos, Angela M. Brown, Richard Peterson, Creighton E. Likes, Dennis K. Watson, and W. Timothy Garvey

Genomic organization and structure of the gene encoding the human uncoupling protein 2 (UCP2) was determined and used to study sequence variations. UCP2 maps on chromosome 11q13 and is a novel member of a gene family comprising three genes (UCP1, UCP2, UCP3) that encode uncouplers of mitochondrial respiration. UCP2 and UCP3 are new candidate genes for obesity and type 2 diabetes (1–3). Human obesity is highly genetically determined as illustrated by heritability estimates ranging from 50 to 85% in twin studies (4). In the mouse, five genes (*ob, db, tub, A^y, fat*) have been described and implicated with various syndromes of monogenic obesity (5). The predominant component of overall energy expenditure in mammals is the basal metabolic rate that reflects nonshivering thermogenesis (4). The uncoupling proteins function to uncouple respiration from oxidative phosphorylation and ATP synthesis and, thus, convert stored fuel into heat. In adult humans, skeletal muscle is the most important source of nonshivering thermogenesis but the molecular mechanism remains unclear. Both UCP2 and UCP3 are expressed in skeletal muscle and, therefore, they have the potential to play an important role in nonshivering thermogenesis and in the pathogenesis of obesity and type 2 diabetes.

To expedite examination of coding regions of the UCP2 gene in diabetic and obese subjects, the genomic organization of the gene was determined and is presented in this report. In addition, a biallelic variant was identified, and allele frequency was determined in populations of Caucasian and African-American subjects. The polymerase chain reaction (PCR) was used to amplify genomic DNA with primers designed from the cDNA with GenBank accession number, U94592.

Human genomic DNA was purified from whole blood and was amplified using the Expand Long Template PCR System as described by the manufacturer (Boehringer Mannheim, Indianapolis, IN). A DNA fragment containing all coding exons was amplified using primers *5′UTRf*: 5′-AGCTTTGAA GAACGGGACAC-3′and *hucp2lr* (5′-CAGAGGTGATCAGG TCAGCA-3′). A 6.5-kb PCR product was generated, and the DNA sequence was determined using additional sequencing primers and an ABI Prism 373 sequencer (Perkin Elmer, Foster City, CA). The sequence was determined bidirectionally with primers designed from the cDNA and from newly generated intronic sequences. All sequence manipulations, assembly, and analyses were performed using GeneWorks software (IntelliGenetics, Mountain View, CA) and the BCM Search Launcher of the Human Genome Center, Baylor College of Medicine. PCR primers, mixtures, and cycling conditions can be found in the on-line appendix at www.diabetes.org/diabetes/appendix.htm.

Determination of the sequence comprising all coding and intronic regions resulted in a continuous DNA sequence of ~4 kb. A schematic of the exon-intron structure is shown in Fig. 1. The "ATG" translation initiator was found to be within what is termed coding exon 1. The entire sequence of the determined coding and noncoding regions is available (GenBank accession number AF019409). No canonical TATA box was identified, and it became apparent that additional noncoding exonic sequences were located upstream. In an attempt to obtain additional 5′ sequence, a 2.5-kb genomic DNA fragment (primer *492r*: 5′-CTCAGAGCCCTTGGTG TAGA-3′ and the Genome Walker kit, Clontech, Palo Alto, CA) was generated. The sequence of the 2.5-kb fragment diverged from that of the cDNA at position –99, indicating the possible presence of an intronic sequence. Accordingly, a splice junction consensus sequence was also identified at this site. Sequence determinations within the 6.5-kb fragment were indicative of an approximate 3-kb intron upstream of the first coding exon, and that cDNA base pairs 5′ of position –99 represented an additional noncoding exon (Fig. 1). Partial sequencing of the 6.5-kb fragment using the upstream primer (*5′UTRf*) yielded an 840-bp sequence that was identical to published cDNA sequence up to position –99, at which point

From the Department of Medicine (G.A., A.M.B., R.P., C.E.L., W.T.G.), Medical University of South Carolina; the Center for Molecular and Structural Biology (D.K.W.), Hollings Cancer Center, Medical University of South Carolina; and the Ralph H. Johnson Veterans Affairs Medical Center (G.A., A.M.B., R.P., C.E.L., W.T.G.), Charleston, South Carolina.

Address correspondence and reprint requests to W. Timothy Garvey, MD, Division of Endocrinology, Diabetes, and Medical Genetics, Medical University of South Carolina, 171 Ashley Ave., Charleston, SC 29425. E-mail: garveywt@musc.edu.

Received for publication 10 September 1997 and accepted in revised form 11 December 1997.

Additional information can be found in an on-line appendix at www.diabetes.org/diabetes/appendix.htm.

PCR, polymerase chain reaction; UCP, uncoupling protein.

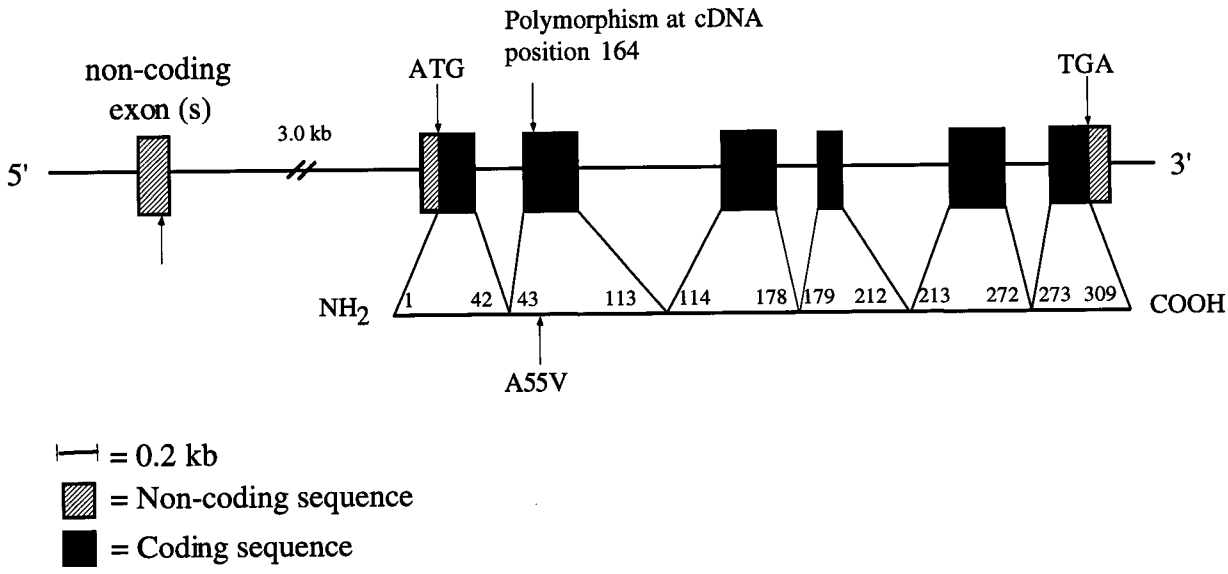

FIG. 1. Schematic representation of the exon-intron structure of the human uncoupling protein 2 gene (UCP2). The sizes of the six coding exons and five intervening introns are indicated and drawn to scale, together with the number of amino acids encoded by each exon below. The interrupted distance between the noncoding exon and the first coding exon represents an estimated 3 kb. The figure assumes one noncoding exon, although genomic sequences were not determined, 5′ of nucleotide-171 based on cDNA nucleotide numbering (accession number U94592), and indicated by a small arrow beneath the noncoding exon.

the sequence diverged (data not shown). At this point of divergence, we observed a consensus sequence for a 5′ splice site ("GG/gtaaga"). By way of comparison, the UCP1 and UCP2 genes have six exons of similar size, each encoding a putative transmembrane domain (3,6). However, the UCP1 gene is less compact (~13 kb) because of expanded introns and does not have a noncoding exon (1).

The UCP2 coding sequence was also examined for the presence of polymorphisms and/or mutations. PCR products (6.5 kb) from patient DNA were generated, and alignments of coding sequences demonstrated identity, except for a mismatch at cDNA nucleotide position 164 (counting from A in the ATG translation initiator) where a cytosine was present in contrast to a thymidine at this position in a previous publication (GenBank accession number U94592). This polymorphism (C164T) results in a conservative substitution of a valine instead of an alanine at codon 55 (A55V). The allele frequency for C164T was established by manual sequencing of DNA fragments amplified from genomic DNA (primers *hucp2lf* and *hucp2lr* and sequenced with primer *492r*) in populations of unrelated Caucasian-Americans (n = 61) and African-Americans (n = 25). However, a quicker and reliable method of restriction digest of a PCR product could also be applied. Specifically, primers *huscp2lf* and *492r* can amplify genomic DNA (PCR conditions can be found in the on-line appendix at www.diabetes.org/diabetes/appendix.htm) and provide a 514-bp fragment. The TT homozygous polymorphism renders the loss of a restriction site (four sites versus five in CC homozygous) for the restriction endonuclease *Cac8 I* (New England Biolabs, Beverly, MA), and the zygocity (CC, CT, TT) can be identified and scored by gel electrophoresis.

Genotype frequency distributions (CC:CT:TT) and allele frequencies were similar in both racial groups with C at cDNA position 164 (62%) occurring more commonly than T (38%) (Table 1). It is unclear whether the conservative substitution of valine for alanine would alter the conformational structure or function of UCP2. This amino acid substitution is predicted to occur in the matrix-oriented extramembranous

TABLE 1
Allele and genotype frequencies for UCP2 gene polymorphism at position 164

	African-Americans	Caucasian-Americans	All subjects (observed)*	All subjects (expected)*
Genotype frequencies†				
CC	10 (40)	31 (51)	41 (48)	33.3 (39)
CT	10 (40)	23 (38)	33 (38)	40.4 (47)
TT	5 (20)	7 (11)	12 (14)	12.3 (14)
Allele frequencies†				
C	60	70	62	
T	40	30	38	

Data are % or n (%). *Expected Hardy-Weinberg distribution, which was established at P = 0.92 compared with observed; †frequency distributions are similar between African-Americans and Caucasian-Americans (NS).

loop between the first and second transmembrane domains based on the hydrophilicity model (3,6). Alanine is conserved at this position in the murine UCP2 gene, but is replaced by tyrosine and leucine in UCP1 and UCP3 genes, respectively. The allele frequencies indicate that C164T is a common gene variant for UCP2 and, thus, has the potential to play a role in large numbers of patients with common polygenic disorders.

The entire coding region of UCP2 was sequenced in five obese and four lean adults as well as in three morbidly obese children (all unrelated), and no sequence variation was detected, other than the C164T polymorphism. However, sequence variation affecting the 5′ untranslated region and upstream regulatory elements have not been addressed. Identification of the common C164T polymorphism will, however, facilitate analyses of the genetic contributions of UCP2 to the development of obesity and diabetes in both Caucasian-Americans and African-Americans. By way of illustration, a common restriction fragment length polymorphism for the UCP1 gene is associated with high weight gain during adult life (7,8) and with resistance to weight loss on a low-calorie diet (9).

ACKNOWLEDGMENTS

This work was supported by grants (W.T.G.) from the National Institutes of Health (DK-47461) and the Department of Veterans Affairs Research Service (Merit Review Program).

We gratefully acknowledge support of the Molecular Core Laboratory of the General Clinical Research Center at the Medical University of South Carolina (M01-RR-1070) for DNA extractions and automated sequencing. We thank Dr. Shichun Bao for providing cDNA samples from human muscle, Dr. Steve Willi for obtaining DNA from three morbidly obese children, Barbara Wojciechowski for statistical analyses, and Pamela Beasley for typing the manuscript.

REFERENCES

1. Cassard A-M, Bouillaud F, Mattei M-G, Hentz E, Raimbault S, Thomas M, Ricquier D: Human uncoupling protein gene: structure, comparison with rat gene and assignment to long arm of chromosome 4. *J Cell Biochem* 43:255–264, 1990
2. Gimeno RE, Dembski M, Weng X, Deng N, Shyjan AM, Gimeno CJ, Iris F, Ellis SJ, Woolf EA, Tartaglia LA: Cloning and characterization of an uncoupling protein homologue: a potential molecular regulator of thermogenesis. *Diabetes* 46:900–906, 1997
3. Vidal-Puig A, Solanes G, Grujic D, Flier JS, Lowell BB: UCP3: an uncoupling protein homologue expressed preferentially and abundantly in skeletal muscle and brown adipose tissue. *Biochem Biophy Res Comm* 235:79–82, 1997
4. Chagnon YC, Peruse L, Bouchard D: Familial aggregation of obesity, candidate genes and quantitative trait loci. *Curr Opin Lipidol* 8:205–211, 1997
5. Dong Z, Guitierrez-Ramos J-C, Coxon A, Mayadas T, Wagner D: A new class of obesity genes encodes leukocyte adhesion receptors. *Proc Natl Acad Sci USA* 94:7526–7530, 1997
6. Klaus S, Casteilla L, Bouillaud F, Ricquier D: The uncoupling protein UCP: a membranous mitochondrial ion carrier exclusively expressed in brown adipose tissue. *Int J Biochem* 23:791–801, 1991
7. Oppert J-M, Vohl M-C, Chagnon M, Dionne FT, Cassard-Doulcier A-M, Ricquier D, Perusse L, Bouchard C: DNA polymorphism in the uncoupling protein (UCP) gene and human body fat. *Int J Obesity* 18:526–531, 1994
8. Clement K, Ruiz J, Cassard-Doulcier A-M, Bouillaud F, Ricquier D, Basdevant A, Guy-Grant B, Froguel P: Additive effect of AG (–3826) variant of the uncoupling protein gene and the trp64arg mutation of the 3-adrenergic receptor gene on weight gain in morbid obesity. *New Engl J Med* 333:352–354, 1995
9. Fumeron F, Durack-Brown I, Betoulle D, Cassard-Doulcier A-M, Tuzet S, Bouillaoud F, Melchior J-C, Ricquier D, Apfelbaum M: Polymorphisms of uncoupling protein (UCP) and 3 adrenoreceptor genes in obese people submitted to a low calorie diet. *Int J Obesity* 20:1051–1054, 1996