

SEQUENCE HETEROGENEITY OF HUMAN LIVER ARGINASE cDNAs AND RESTRICTION FRAGMENT LENGTH POLYMORPHISM OF THE GENE LOCUS

Yougo HARAGUCHI,^{1,2} Masaki TAKIGUCHI,¹ Ichiro MATSUDA,²
and Masataka MORI¹

¹*Institute for Medical Genetics and* ²*Department of Pediatrics,*
Kumamoto University Medical School, Kumamoto 862, Japan

Summary Arginase is an enzyme involved in the urea cycle and its deficiency results in argininemia, an autosomal recessive disorder accompanied by hyperammonemia. We isolated two overlapping cDNA clones and determined the entire nucleotide sequence (Haraguchi *et al.* 1987. *Proc. Natl. Acad. Sci. USA* 84: 412-415). Sequence analysis of eight cDNA clones revealed nucleotide variations at five positions in the protein coding region, and all these variations were accompanied by amino acid substitutions (Met to Lys at amino acid residue 1, Gln to Glu at residue 86, Ser to Phe at residue 109, Leu to Gln at residue 206, and Val to Ala at residue 233). These results suggest that human liver arginase is polymorphic in amino acid sequence, although some of the variations may be due to a transcription error or to cloning artifacts. Two restriction fragment length polymorphisms were identified at the arginase gene locus, using restriction endonucleases *PvuII* and *HincII*. These sequence polymorphisms and restriction fragment length polymorphisms may serve as linkage markers for arginase deficiency.

INTRODUCTION

Arginase is the last enzyme functional in the urea cycle and catalyzes the conversion of arginine to urea and ornithine in the liver of ureotelic animals. Arginase from rat (Schimke, 1962; Hirsch-Kolb and Greenberg, 1968; Tarrab *et al.*, 1974), rabbit (Vielle-Breitburd and Orth, 1972) and human (Beruter *et al.*, 1978; Brusdeilins *et al.*, 1985) livers has been purified and was found to consist of three identical subunits (Penninckx *et al.*, 1974) of about 35,000 daltons. Arginase isozyme(s) that differs from the liver enzyme in catalytic, molecular and immunological properties, is present in the kidney, small intestine and lactating mammary gland (Glass and Knox, 1973; Herzfeld and Raper, 1976; Skrzypek-Osiecka *et al.*, 1983; Spector

Received January 14, 1988; Accepted January 28, 1988

et al., 1983). A deficiency in human liver arginase results in argininemia, an inherited autosomal recessive disorder accompanied by hyperammonemia (Walser, 1983). We isolated cDNA clones for rat (Kawamoto *et al.*, 1987) and human (Haraguchi *et al.*, 1987) liver arginase and determined the nucleotide sequences. Sparkes *et al.* (1986) also isolated human liver arginase cDNA and assigned the gene to chromosome band 6q23. During isolation of human arginase cDNA clones, we found several distinct groups of cDNAs which differed in restriction maps, therefore, we determined the entire sequences of eight independent clones. We now report nucleotide and amino acid sequence heterogeneity of human liver arginase cDNA and protein. Two restriction fragment length polymorphisms (RFLPs) of the gene locus are also reported.

MATERIALS AND METHODS

Isolation of human liver arginase cDNA clones. An oligo(dT)-primed human liver cDNA library in λ gt11 and a random primer library in λ gt11 were screened by plaque hybridization, as described (Haraguchi *et al.*, 1987). The cDNA libraries were kindly provided by G.A. Ricca (Meloy Laboratories, Springfield, Va., USA) and by Y. Ebina (Kumamoto University) and J. Ou (University of California, San Francisco, Calif., USA), respectively. Each library was prepared from the hepatic mRNA of one individual. Plaques of positive clones were purified, and *Eco*RI-excised cDNA inserts were subcloned into a plasmid vector pUC9 and characterized by restriction endonuclease mapping.

Nucleotide sequence determination. The pUC9 subclones were directly subjected to nucleotide sequencing by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977), using an alkali-denatured plasmid as a template (Hattori and Sakaki, 1986) and a synthetic oligonucleotide as a primer. Oligonucleotide primers were synthesized using a DNA synthesizer (Model 381A, Applied Biosystems, Foster City, Calif., USA).

Southern blot analysis. DNA was prepared from peripheral blood lymphocytes, by the established procedure (Bell *et al.*, 1981). DNA (10 μ g) was digested for 18 hr with restriction endonucleases *Pvu*II, *Taq*I, or *Hinc*II (Takara Shuzo, Kyoto, or Toyobo Biochemicals, Tokyo; 50 units each) under the conditions recommended by the supplier. The digests were separated by agarose gel electrophoresis (0.7–0.8% agarose, 14–16 hr at 1.9 V/cm), and transferred to nitrocellulose filters. Hybridization was carried out according to Maniatis *et al.* (1982) using nick-translated, 32 P-labeled cDNA (full-length, 1×10^8 dpm/ μ g DNA) as a probe. The filters were then washed and exposed for several days at -70°C to a Fuji RX film with an intensifying screen.

RESULTS

Sequence heterogeneity

About 1.2×10^5 clones of an oligo(dT)-primed human liver cDNA library in λ gt11 and 5×10^5 clones of a random primer library in λ gt11 were screened with the 1,350-base pair (bp) *EcoRI-PstI* fragment of the rat liver arginase cDNA plasmid pARGr-2 (Kawamoto *et al.*, 1986) as a probe. The cDNA inserts of four positive phage clones from each library were subcloned into a plasmid vector pUC9 and designated phARG 1, 2, 3, 6, 101, 103, 109 and 112 (Fig. 1). Among these plasmids, the plasmid phARG 6 was identified as an arginase cDNA clone, because the bacterial clone carrying this plasmid had an arginase activity whereas the bacteria carry-

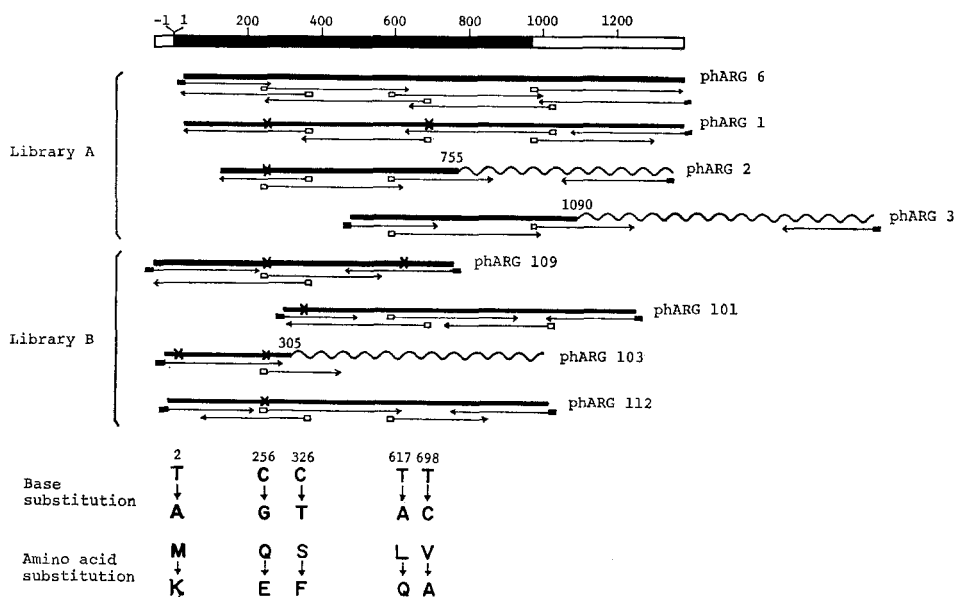


Fig. 1. Sequence polymorphism in human liver arginase cDNA and the strategy for nucleotide sequence determination. The scale at the top indicates the nucleotide position relative to the protein initiator codon beginning at position 1. The protein coding region is indicated by solid box. The arrows beneath each cDNA indicate the direction and extent of sequence determination. Wavy lines show sequences completely different from the arginase cDNA sequence (see text); the numbers above the junctions are nucleotide positions. The positions of the base substitutions are marked by x, and these substitutions and the resulting amino acid substitutions are shown at the bottom. The numbers above the base substitutions indicate the nucleotide positions. Libraries A and B show an oligo(dT)-primed human liver cDNA library in λ gt11 and a random primer library in λ gt11, respectively.

ing no plasmid or the cloning vector pUC9 had no detectable activity (Haraguchi *et al.*, 1987). The entire nucleotide sequence of phARG 6 (which lacked the NH₂-terminal eight amino acid residues) and the 5' region sequence of phARG 109 which extended further into the 5' direction, were determined (Haraguchi *et al.*, 1987). Sequencing of the remaining portion of phARG 109 revealed two base substitutions between these two cDNA clones (Fig. 1). C at nucleotide position 256 and T at position 617 in phARG 6 were replaced by G and A, respectively. Both base substitutions were accompanied by amino acid substitutions, Gln to Glu and Leu to Gln, respectively (Fig. 1). Profound variations were seen in restriction maps between the eight clones. We then examined almost all of the nucleotide sequences of these clones and the results are summarized in Fig. 1. phARG 2, phARG 3 and phARG 103 were composed of human liver arginase cDNAs and distinct sequences quite different from that of arginase cDNA. Among these sequences, one (in phARG 103) was found to be the 5' portion of intron 3 which was not spliced out from the mRNA precursor for human liver arginase (unpublished results), whereas the other

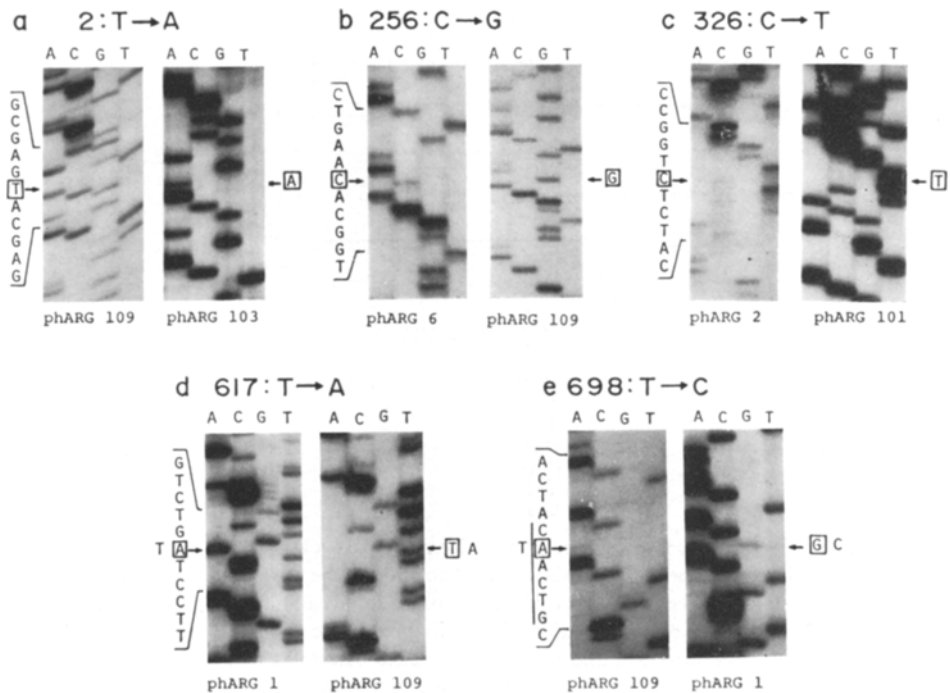


Fig. 2. Base substitutions in human liver arginase cDNA. Relevant portions of sequencing gels are shown. The nucleotide positions and the base substitutions are shown at the top of each panel. Arrows indicate the positions of nucleotide alterations and the altered nucleotides are boxed. d and e show the antisense-strand nucleotide sequences; the sense-strand nucleotides are shown outside. *HincII* site in e is shown by the vertical sideline.

two (in phARG 2 and phARG 3) might be distinct cDNAs of unknown nature (see DISCUSSION). Therefore, the considerable variation in the restriction maps of the cDNAs is due mainly to the presence of these arginase cDNA-unrelated sequences.

Among the eight arginase cDNA sequences, nucleotide variations were found at five positions in the protein-coding region, including the two positions mentioned above. Sequencing gels showing the base substitutions are shown in Fig. 2. T to C substitution at position 698 resulted in the disappearance of the *HincII* site. All these nucleotide variations were accompanied by amino acid substitutions (Met to Lys at residue 1, Ser to Phe at residue 109, Val to Ala at residue 233, and the two variations mentioned above). At all these variable positions, minor variant nucleotides were found only in one clone. The corresponding nucleotides determined from human liver genomic clones (unpublished results) were identical with the major nucleotide forms.

The positions of the amino acid substitutions are shown along with the sequence comparison of human liver, rat liver and yeast arginases (Fig. 3). The Met-Lys variation at amino acid residue 1 resulted in a loss of the initiator methionine. With respect to Gln-Glu variation at residue 86, Glu is the major form and is conserved in the sequence in the rat. For Ser-Phe variation at residue 109, Ser is the major form and is conserved among the three species. For the Leu-Gln variation at residue

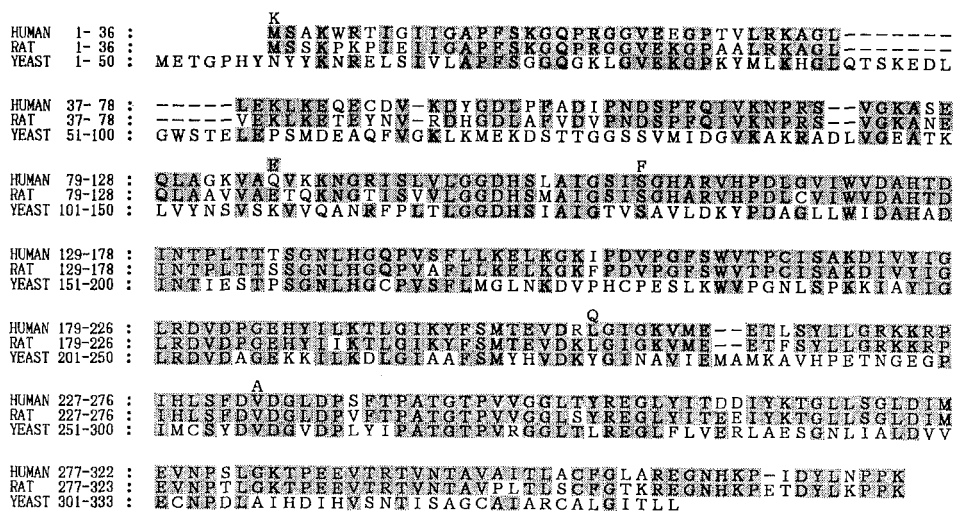


Fig. 3. Positions of the amino acid substitutions in human liver arginase. The amino acid sequence of human liver arginase and that of rat liver arginase are from our reported results (Haraguchi *et al.*, 1987; Kawamoto *et al.*, 1987), and that of yeast enzyme is from Sumrada and Cooper (1984). Gaps were introduced to increase the similarity and matching amino acids are shaded. Amino acid substitutions are shown above the human sequence.

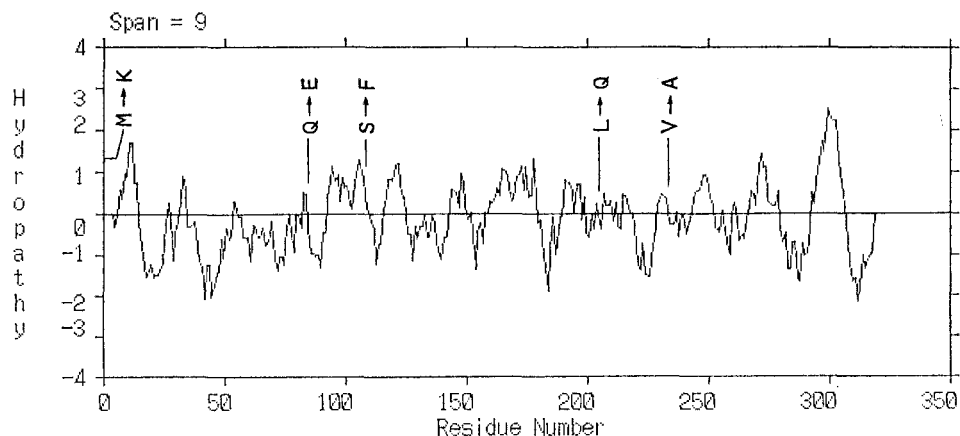


Fig. 4. Hydropathy profile of human liver arginase and positions of amino acid substitutions. Hydropathy profile was analyzed according to Kyte and Doolittle (1982). The amino acid substitutions are shown above the hydropathy profile.

206, Leu is the major form and is conserved in the sequence in the rat. For the Val-Ala variation at position 233, Val is the major form and is conserved among the three species.

Hydropathy profiles of human liver arginase and positions of the amino acid substitutions are shown in Fig. 4. The arginase has a number of relatively short hydrophobic and hydrophilic regions. The residue 1 of Met-Lys variation is located at a hydrophilic region (the hydropathy profile of this region is not shown). The residues of the three variations (residues 86, 109 and 233) are situated at regions of hydropathy value of near zero, whereas that of the other variation (residue 206) is situated at a hydrophilic region.

RFLPs of the arginase gene locus

DNAs from healthy Japanese individuals were digested with restriction endonucleases *PvuII*, *TaqI* or *HincII*, and the digests were subjected to Southern blot analysis using human liver arginase cDNA as a probe. No RFLP was found with *TaqI* (data not shown). On the other hand, an RFLP was found with *PvuII* (Fig. 5A). The 2.8 kilobase (kb) fragment was present in only 2 of the 14 individuals. This 2.8 kb fragment appeared to be allelic to 1.5 kb and 1.3 kb fragments. Thus, individuals 1 and 2 were heterozygous for the 2.8 kb allele and the 1.5 kb and 1.3 kb allele, whereas the other individuals were homozygous for the 1.5 plus 1.3 kb allele. Frequency of the 2.8 kb allele was 2/28 or about 7%. Another RFLP was found with *HincII*. The 1.2 kb fragment which is apparently allelic to 1.5 kb fragment, was present only in individuals 4 and 11 but not in others. Thus, the former two individuals were heterozygous for the 1.5 kb and 1.2 kb allele whereas the other individuals were homozygous for the 1.5 kb allele. Frequency of the 1.2 kb allele

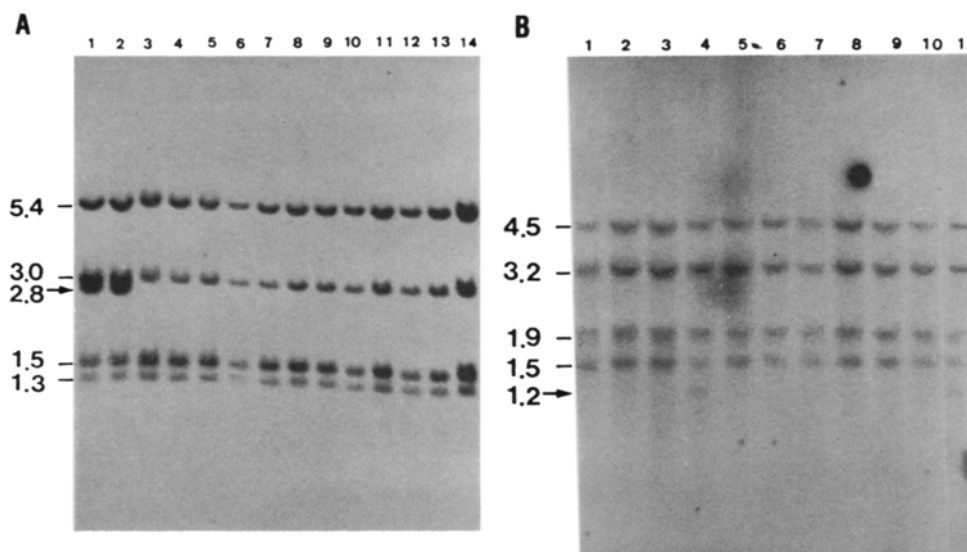


Fig. 5. Southern blot analysis of DNA from a healthy Japanese subject hybridized with human liver arginase cDNA. Ten micrograms DNA from healthy 14 (A) or 11 (B) individuals were digested with 50 units of *Pvu*II (A) or *Hinc*II (B) and the digests were separated by electrophoresis in 0.7% (A) or 0.8% (B) agarose gel. The separated DNAs were transferred to nitrocellulose filters and hybridized with the nick-translated, ³²P-labeled cDNA. Band sizes are given in kb. Polymorphic sites are illustrated by arrows. Individuals 1, 2, 5, 6 and 7 in (A) are identical with individuals 11, 4, 1, 2 and 3 in (B), respectively.

was 2/22 or 9%. To be noted here is that individuals 1 and 2 in Fig. 5A are the same as individuals 4 and 11 in Fig. 5B. This indicates that the two RFLPs are closely linked.

DISCUSSION

Among eight cDNA clones for human liver-type arginase, three clones (phARG 2, phARG 3 and phARG 103) contained sequences different from that of arginase cDNA. Sequence analysis of genomic clones (unpublished results) revealed that the relevant sequence in phARG 103 completely coincided with that of the 5' portion of intron 3. Thus, it is clear that the clone phARG 103 arose from an arginase mRNA species containing the unspliced intron 3. On the other hand, the relevant sequences in phARG 2 and phARG 3 are not from unspliced introns. It is likely that the two cDNA clones were formed by fusion of arginase cDNAs and distinct cDNAs, during the cloning procedures. The nature of the distinct sequences is unknown.

In addition to these arginase cDNA-unrelated sequences, five base alterations were found in the protein-coding region among the eight cDNA clones and all these base alterations were accompanied by amino acid alterations. Among the five alterations, the T-A alteration at nucleotide position 2 and C-T alteration at position 326 may be due to cloning artifacts because they were situated near 5' termini of the clones and because the substitution at position 2 resulted in a loss of the initiation methionine codon. If these two variations are omitted, one cDNA library (oligo(dT)-primed, Library A in Fig. 1) contained three distinct cDNA species and the other library (random primer-primed, Library B in Fig. 1) contained two distinct cDNAs. Each library was constructed from hepatic mRNA of one individual. Multiple cDNA species may arise from the two alleles, multiple gene loci, RNA polymerase errors or cloning artifacts. Genomic Southern blotting and genomic cloning revealed a single copy gene for human liver arginase (unpublished results). It is very unlikely that all these variations are due to RNA polymerase errors or cloning artifacts, although one of the three distinct cDNAs in Library A may be due to one of these errors. All these results strongly suggest that the two distinct cDNAs in each library are allelic and that the human liver arginase is highly polymorphic in amino acid sequence. These sequence polymorphisms can be checked using sequence-specific oligonucleotide probes.

Two RFLPs of the human liver arginase gene were found with the enzyme *PvuII* and *HincII*. The frequencies of the minor components of these two RFLPs are 7 and 9%, respectively. These two RFLPs are closely linked. Kidd *et al.* (1986) noted a similar *PvuII* RFLP of the arginase gene in Caucasians. Frequency of the minor 2.8 kb allele was 14% (91 individuals). Thus, no marked difference was observed in allele frequency between Japanese and Caucasian populations. On the other hand, in contrast to our findings, they detected no RFLP with *HincII*, when tested on 13 individuals. The reason of this discrepancy has yet to be determined. In any case, these RFLPs can be used in the DNA diagnosis of argininemia, only in limited cases. Other RFLPs are now being searched for using other restriction endonucleases and, as probes, genomic fragments. If the sequence polymorphism mentioned above is valid, then the DNA diagnosis using sequence-specific oligonucleotides will be feasible.

Acknowledgments We thank Dr. G.A. Ricca for providing the oligo(dT)-primed human liver cDNA library, Drs. Y. Ebina and J. Ou for providing the random primer library, Drs. E. Araki (Kumamoto University) and Y. Ebina for providing peripheral blood DNAs, and M. Ohara (Kyushu University, Japan) for comments on the manuscript. This work was supported in part by Grants-in-Aid 61870019 and 62109008 from the Ministry of Education, Science and Culture of Japan, and a grant from the Naito Foundation.

REFERENCES

- Bell, G.I., Karam, J.H., and Rutter, W.I. 1981. Polymorphic DNA region adjacent to the 5' end of the human insulin gene. *Proc. Natl. Acad. Sci. USA* **78**: 5759-5763.

- Beruter, J., Colombo, J.P., and Bachmann, C. 1978. Purification and properties of arginase from human liver and erythrocytes. *Biochem. J.* **175**: 449-454.
- Brusdeilins, M., Kuhner, R., and Schumacher, K. 1985. Purification, affinity to anti-human arginase immunoglobulin-Sepharose 4B and subunit molecular weights of mammalian arginases. *Biochim. Biophys. Acta* **840**: 79-90.
- Glass, R.D. and Knox, W.E. 1973. Arginase isozymes of rat mammary gland, liver, and other tissues. *J. Biol. Chem.* **248**: 5785-5789.
- Haraguchi, Y., Takiguchi, M., Amaya, Y., Kawamoto, S., Matsuda, I., and Mori, M. 1987. Molecular cloning and nucleotide sequence of cDNA for human liver arginase. *Proc. Natl. Acad. Sci. USA* **84**: 412-415.
- Hattori, M. and Sakaki, Y. 1986. Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* **152**: 232-238.
- Herzfeld, A. and Raper, S.M. 1976. The heterogeneity of arginases in rat tissues. *Biochem. J.* **153**: 469-478.
- Hirsch-Kolb, H. and Greenberg, D.M. 1968. Molecular characteristics of rat liver arginase. *J. Biol. Chem.* **243**: 6123-6129.
- Kawamoto, S., Amaya, Y., Oda, T., Kuzumi, T., Saheki, T., Kimura, S., and Mori, M. 1986. Cloning and expression in *Escherichia coli* of cDNA for arginase of rat liver. *Biochem. Biophys. Res. Commun.* **136**: 955-961.
- Kawamoto, S., Amaya, Y., Murakami, K., Tokunaga, F., Iwanaga, S., Kobayashi, K., Saheki, T., Kimura, S., and Mori, M. 1987. Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver arginase. *J. Biol. Chem.* **262**: 6280-6283.
- Kidd, J.R., Dizikes, G.J., Grody, W.W., Cederbaum, S.D., and Kidd, K.K. 1986. A *PvuII* RFLP for the human liver arginase (ARG) gene. *Nucleic Acids Res.* **14**: 2544.
- Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. 1982. *Molecular Cloning*, Cold Spring Harbor Laboratory, New York.
- Penninckx, M., Simon, J.P., and Wiame, J.M. 1974. Interaction between arginase and L-ornithine carbamoyltransferase in *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **49**: 429-442.
- Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.
- Schimke, R.T. 1962. Adaptive characteristics of urea cycle enzymes in the rat. *J. Biol. Chem.* **237**: 459-468.
- Skrzypek-Osiecka, Robin, Y., and Poremska, Z. 1983. Purification of rat kidney arginase A1 and A4 and their subcellular distribution. *Acta Biochim. Pol.* **30**: 83-92.
- Sparkes, R.S., Dizikes, G.J., Klisak, I., Grody, W.W., Mohandas, T., Heinzmann, C., Zollman, S., Lusic, A.J., and Cederbaum, S.D. 1986. The gene for human liver arginase (ARG1) is assigned to chromosome band 6q23. *Am. J. Hum. Genet.* **39**: 186-193.
- Spector, E.B., Rice, S.C.H., and Cederbaum, S.D. 1983. Immunologic studies of arginase in tissues of normal human adult and arginase-deficient patients. *Pediatr. Res.* **17**: 941-944.
- Sumrada, R.A. and Cooper, T.G. 1984. Nucleotide sequence of the *Saccharomyces cerevisiae* arginase gene (CAR1) and its transcription under various physiological conditions. *J. Bacteriol.* **160**: 1078-1087.
- Tarrab, R., Rodriguez, J., Huitron, C., Palacios, R., and Soberon, G. 1974. Molecular forms of rat-liver arginase. Isolation and characterization. *Eur. J. Biochem.* **49**: 457-468.
- Vielle-Breitburd, F. and Orth, G. 1972. Rabbit liver L-arginase. *J. Biol. Chem.* **247**: 1227-1235.
- Walser, M. 1983. Urea cycle disorders and other hereditary hyperammonemic syndromes. In *The Metabolic Basis of Inherited Disease*, Stanbury, J.B., Wyngaarden, J.B., Fredrickson, D.S., Goldstein, J.L., and Brown, M.S. eds., McGraw Hill, New York, pp. 402-438.