European Journal of Human Genetics (2000) 8, 725–729

www.nature.com/ejhg

SHORT REPORT

Genomic structure and transcript variants of the human methylenetetrahydrofolate reductase gene

Anja Homberger, Michael Linnebank, Carsten Winter, Holger Willenbring, Thorsten Marquardt, Erik Harms and Hans Georg Koch

Universität Münster, Klinik und Poliklinik für Kinderheilkunde, Albert-Schweitzer-Strasse 33, 48129 Münster, Germany

The human 5,10-methylenetetrahydrofolate reductase (MTHFR) represents a major enzyme in the folate-dependent regulation of methionine and homocysteine concentrations. Different MTHFR mutations lead either to severe homocystinuria as a multisystem disorder or to moderate hyperhomocysteinaemia, which is a common risk factor for disorders ranging from cardiovasculopathy to spina bifida. The N-terminal part of the human *MTHFR* gene is incompletely characterised. We report the completed genomic structure of this gene including three novel exonic sequences on the basis of a 5'-RACE and a 4.2 kb cloned fragment of human genomic DNA. We demonstrate the existence of four MTHFR transcripts differing in their first exons. The diversity of transcripts is due to alternative transcription initiation and alternative splicing. Three putative polypeptides of 657, 698, and 680 amino acids are encoded. The novel genomic sequence described here includes putative promoter regions as suggested by the presence of regions homologue to binding sites for SP1, AP1, AP2, CAAT or GC boxes. Furthermore, we provide evidence that there are no TATA-box elements to regulate the human *MTHFR* gene. The results of our study render the full-length characterisation of affected alleles in severe homocystinuria and moderate hyperhomocysteinaemia due to MTHFR deficiency and provide a basis for investigating the regulation of the human *MTHFR* gene. *European Journal of Human Genetics* (2000) **8**, 725–729.

Keywords: methylenetetrahydrofolate reductase; MTHFR; promoter; transcripts; genomic structure

Introduction

The human 5,10-methylenetetrahydrofolate reductase (MTHFR Enzyme Commission 1.5.1.20) is a key enzyme of folate metabolism and is essential for the regulation of methionine and homocysteine concentrations. The homodimer MTHFR catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which is the carbon donor in the remethylation of homocysteine to methionine. In human tissues a major polypeptide of about 77 kDa was found by western blot analysis, whereas an additional MTHFR polypeptide of about 70 kDa was detected only in human liver tissue.[1] The different sizes of the putative isoforms are so far of unexplained origin.

Goyette *et al* described the partial genomic organisation of the human *MTHFR* gene, which had been mapped to chromosome 1p36.3, and identified 11 exons. However, the N-terminal part including the promoter region remained unknown.[2,3]

In the small group of patients suffering from severe homocystinuria (MIM 236250) at least 18 different, mostly private mutations of the *MTHFR* gene have been identified,[2,4–6], whereas the frequent mutations 677C > T and 1298A > C are linked to moderate hyperhomocysteinaemia.[1,7]

The aim of this study was to complete the coding sequence as well as the genomic structure of the *MTHFR* gene in order to provide a basis for full-length characterisations of rare and common variants as well as for investigations on the genetic regulation of human MTHFR.

Correspondence: Priv. Doz. Dr HG Koch, Universität Münster, Klinik und Poliklinik für Kinderheilkunde, Albert-Schweitzer-Str. 33, D-48129 Münster, Germany. Tel: +49 251 8356494; Fax: +49 251 8356085; E-mail: kochha@uni-muenster.de
Received 23 January 2000; revised 8 May 2000; accepted 18 May 2000

Material and methods

Total RNA was prepared from cultured fibroblasts using the RNeasy system (Qiagen, Hilden, Germany). The mRNA fraction was prepared using the mRNA Isolation Kit (Boehringer Mannheim, Mannheim, Germany) and served as a template for 5'-RACE subsequently using the RACE-primers 5'-GTTCCAGGGCAGGCAAGT, 5'-GCCAGCTCGATGCCA-TAGTTGC, and 5'-TGAGATGAGATTGACAGCTC, together with the 5'-RACE kit essentially according to the manufacturer's instructions (Boehringer Mannheim). Products generated by 5'-RACE were purified (QiaQuick gel extraction kit, Qiagen), directly cloned into pCR2.1 (Invitrogen, De Shelp, The Netherlands), and transferred to DNA sequence analysis on the LI-COR 4000L sequencing device according to the manufacturer's instructions (LICOR, Lincoln, NB, USA).

For verification of the natural occurrence of transcript variants we performed transcript specific RT-PCR reactions using specific sense primers (MTHFR1: 5-AGTCGTAGGCT-TAGTATCCC, pos. –2261 to –2242; MTHFR2: 5'-CACTCTGGGCCTGAGCTGAC, pos. –410 to –391; MTHFR3: 5'-GCTTGAACCTGCCACTCAGG, pos. –84 to –65) and a common reverse PCR primer (5'-TGAGATGAGATTGA-CAGCTC, pos. 234 to 215) which was tailed with a sequence complementary to the M13 reverse primer to allow direct cycle sequencing of PCR products. We used cDNA generated from human cultured fibroblasts as a template for PCR.

Genomic clones harbouring the N-terminal region of the human *MTHFR* gene were isolated from a human γ ZAP Express® library following standard procedures (Stratagene, Heidelberg, Germany). A 0.6 kb DNA probe homologue to the 5'-region of the MTHFR transcripts was digoxigenin labelled during PCR with the primer pair 5'-CACTCTGGGCCT-GAGCTGAC/5'-GGAAGAATTCCAGGGAGAAC (Boehringer Mannheim) and used for isolation of a genomic clone. Overlapping *Eco*RI and *Pst*I subfragments were sequenced on both strands.

Results

We applied the 5'-RACE technology in order to characterise the N-terminal region of the human *MTHFR* gene. Analysing mRNA derived from human cultured fibroblasts we identified a transcript providing novel sequence information on the *MTHFR* gene (AJ237672), which we used for screening a genomic library. We isolated a genomic clone containing a 4.2 kb fragment, which was sequenced entirely (AJ249275). When compared with the already characterised part of the human *MTHFR* gene there was 3.8 kb of novel N-terminal sequence information. Furthermore, screening of genomic databases using the BLAST software (http://vega.igh.cnrs.fr/bin/blast-guess.cgi) revealed matching to a putative human chloride channel gene *CLCN6* (AF009247). It was localised on the opposite strand approximately 0.5 kb apart from the exonic MTHFR sequences which were identified in this study. Two TATA-box motifs were found in the *CLCN6* gene, which are possibly involved in the regulation of this gene (exact nucleotide positions are available, AJ249275).

The 5'-RACE experiments revealed further transcripts identical in the 3'-part, but different in the 5'-sequences. Comparison of transcriptional and genomic sequences allowed to deduce the structure of the so far unknown part of the human *MTHFR* gene (Figure 1, Figure 2). In total, we identified three different transcripts (MTHFR1, 2, and 3), which differed in their first exons (Figure 3). Additionally, transcript MTHFR 1 was found in two variants which were discriminated by the absence or presence of three nucleotides (g.-1 to 2) due to alternative involvement of two adjoined consensus splice signals without affecting the deduced polypeptide sequences (see AJ249275). Each of the three major MTHFR transcripts 1, 2, and 3 had an individual first ATG codon

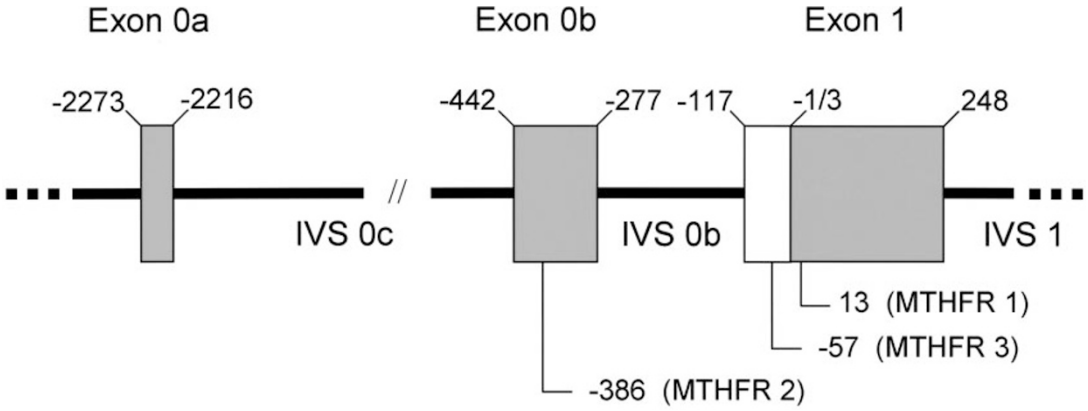

Figure 1 Genomic structure. The positions of genomic sequences of the human *MTHFR* gene used as exons are depicted by shaded boxes. The putative translation initiation sites of the identified transcripts MTHFR 1, 2, and 3 are shown underneath. The sequence spanning nucleotide positions g.–177 to g.–2 (white box) is used as exonic sequence in the transcript MTHFR 3, but is spliced out of MTHFR 1 and 2. The two variants of MTHFR 1 are distinguished due to different usage of the two splice-acceptor sites within exon 1 at nucleotide position 1 or 3.

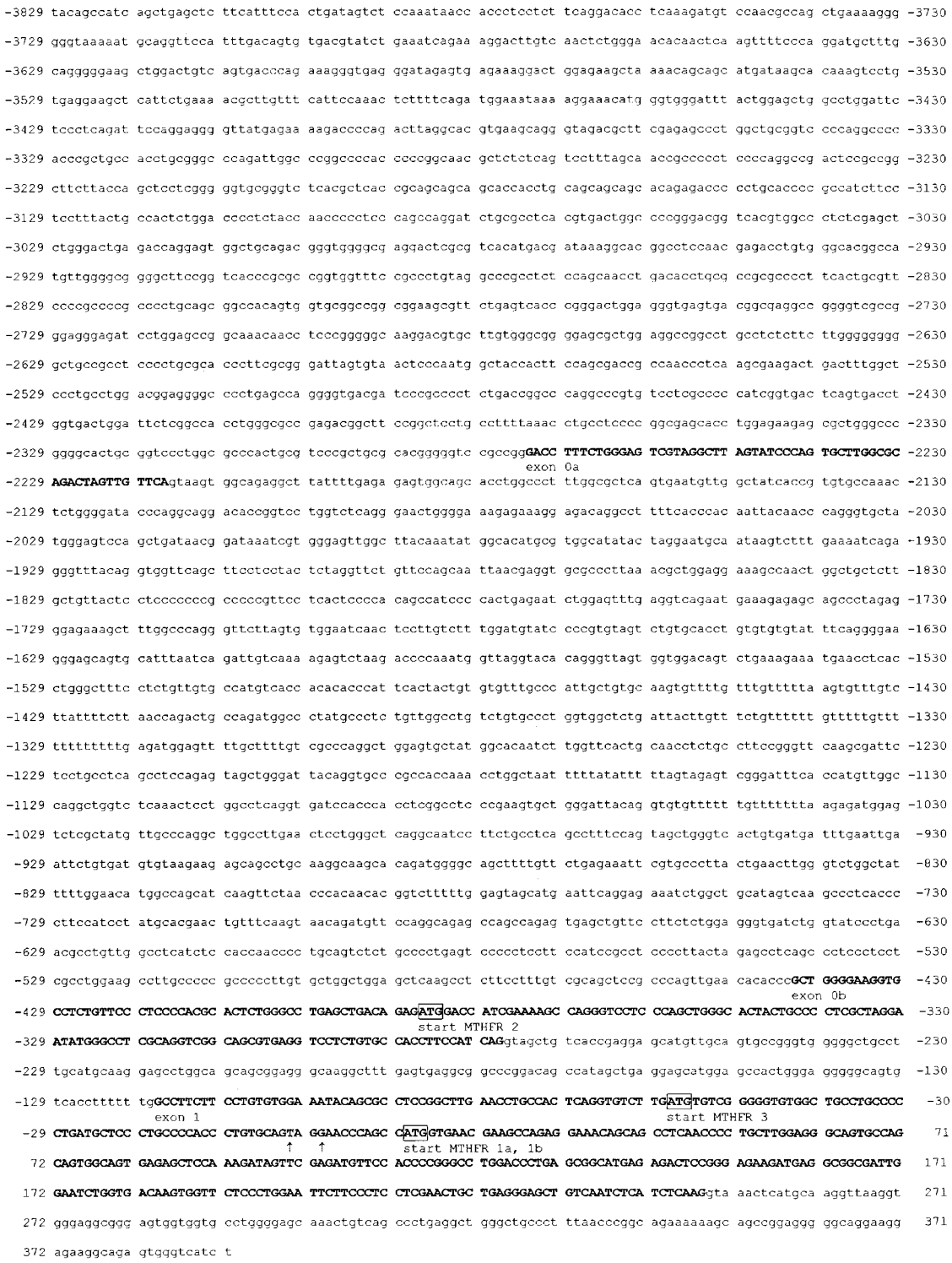

Figure 2 Sequence of genomic DNA. DNA sequence derived from a cloned fragment containing the N-terminal part of the human MTHFR gene. Exonic sequences are given in bold capitals. ATG start codons of the transcripts MTHFR 1, 2, and 3 are depicted in frames. TATA box motifs on the opposite strand possibly involved in the regulation of an adjoined putative chloride channel gene were identified at nucleotide positions –2394 to –2380 and –1180 to –1171, respectively. The borders of the two splice-acceptor sites within exon 1 upstream nucleotides –1 or 3, respectively, are indicated by arrows.

predicting different open reading frames of 1971, 2094, and 2040 bp, possibly encoding polypeptides of 657, 698, and 680 amino acid residues, respectively. The natural occurrence of each transcript was verified by PCR amplification of cDNA derived from cultured fibroblasts. For each of the transcripts MTHFR1, 2, and 3, PCR products of the expected lengths could be detected and were proved for their specificity by sequencing. Analysis of the genomic sequence provided in this paper with the MOTIF software (http://www.motif.genome.ad.jp/) demonstrated the absence of any TATA-boxes. However, there are several consensus sequences for potential binding sites of transcription factors, eg SP1, AP1, AP2, CAAT or GC-boxes (data not shown).

Discussion

MTHFR is a critical enzyme for the remethylation of homocysteine. A large body of controversial data exists regarding the epidemiological impact of frequent MTHFR variants as risk factors for vasculopathy and neural tube defects.[8,9] The regulation of this gene could not yet be elucidated, since the regulatory domains remained obscure. The novel sequence information in this study provide the missing data to investigate the regulation of the gene and might be helpful for the complete molecular characterisation of MTHFR-deficient patients. Furthermore, the sequence conjunction of the *MTHFR* gene with the *CLCN6* gene harbours detailed information on the physical position of both genes, since the human *MTHFR* gene was mapped to chromosome 1p36.3.[2]

The identification of different transcripts allowed deduction of the genomic structure of the so far unknown part of the human *MTHFR* gene and points to a complex regulation on the transcriptional level. The existence of different first exons is most likely due to alternative transcription initiation. The individual first ATG codons of all identified transcripts are in frame. The recently reported furthest aminoterminal ATG is identical with the putative start codon in *MTHFR 1*[3] and fits best to the Kozak consensus sequence by means of translation initiation (data not shown). *MTHFR 2* and *3*, despite containing additional aminoterminal sequences including individual ATGs (Figures 1 and 3), might also use this ATG as a start codon. Our data do not provide evidence of the biological relevance of transcriptional variability, but polypeptides of distinct sizes of approximately 70

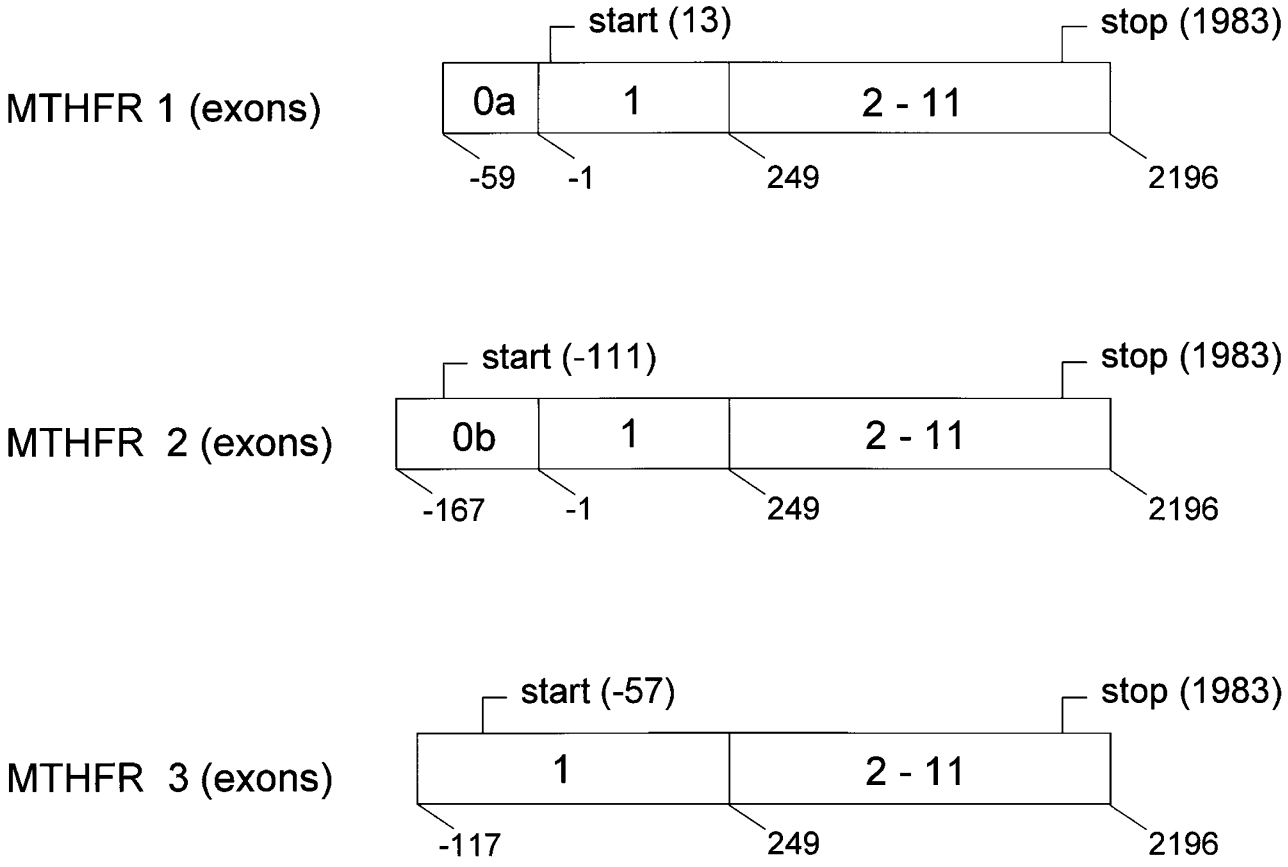

Figure 3 Human MTHFR transcript variants. The different transcripts (MTHFR 1, 2, and 3) of the human *MTHFR* gene are determined by specific usage of the respective first exons. The numbers underneath indicate the 5'-border of each exon or the end of the transcript. The transcripts reveal individual open reading frames with unique start sites but constant stop codons. The indicated nucleotide positions refer to Goyette *et al*[3].

and 77 kD could be detected physiologically depending on the type of tissues analysed.[1] The molecular weights of the putative MTHFR polypeptides deduced from the different transcripts were calculated to be 74.5, 78.9, and 76.9 kD, respectively, employing the SAPS (statistical analysis of protein sequences) software (http://www.isrec.isb-sib.ch/software/SAPS—form.html). We cannot yet decide whether the data derived from protein analysis correspond to the polypeptide sizes deduced from the transcriptional variability presented here.

This study describes the putative promoter region of the human *MTHFR* gene containing numerous consensus sequences for transcription factor binding sites, but lacking a TATA-box. There are striking similarities to other genes involved in regulation of homocysteine metabolism, cystathionine β-synthase (*CBS*), methionine synthase (*MTR*), methionine synthase reductase (*MTRR*), all sharing consensus sequences for transcription factor binding sites as mentioned above. Furthermore, TATA-elements are not present in any of this genes, and different transcription and/or translation and/or alternative splicing events were assumed to be involved in their regulation.[10–13]

In conclusion, this study reports the so far unknown part of the human *MTHFR* gene including the putative promoter region and the genomic organisation. The regulation of *MTHFR* may now be investigated to obtain new insights into the methionine and homocysteine metabolism for a better understanding of dysfunctions associated with common moderate hyperhomocysteinaemia or rare severe MTHFR deficiency.

Acknowledgements

This work was supported by the Interdisziplinäres Zentrum für Klinische Forschung of the University of Münster, Germany (IZKF-B2). The technical assistance of I Du Chesne is gratefully acknowledged. Anja Homberger and Michael Linnebank contributed equally to this work. EMBL database accession numbers AF237672 and AF249275.

References

1 Frosst P, Blom HJ, Milos R *et al*: A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet* 1995; **10**: 111–113.
2 Goyette P, Summer JS, Milos R *et al*: Human methylenetetrahydrofolate reductase: isolation of cDNA, mapping and mutation identification. *Nat Genet* 1994; **7**: 195–200.
3 Goyette P, Pai A, Milos R *et al*: Gene structure of human and mouse methylenetetrahydrofolate reductase (MTHFR). *Mamm Genome* 1998; **9**: 652–656.
4 Goyette P, Frosst P, Rosenblatt DS, Rozen R: Seven novel mutations in the methylenetetrahydrofolate reductase gene and genotype/phenotype correlations in severe methylenetetrahydrofolate reductase deficiency. *Am J Hum Genet* 1995; **56**: 1052–1059.
5 Goyette P, Christensen B, Rosenblatt DS, Rozen R: Severe and mild mutations in *cis* for the methylenetetrahydrofolate reductase (MTHFR) gene, and description of five novel mutations in MTHFR. *Am J Hum Genet* 1996; **59**: 1268–1275.
6 Kluijtmans LA, Wendel U, Stevens EM, van den Heuvel LP, Trijbels FJ, Blom HJ: Identification of four novel mutations in severe methylenetetrahydrofolate reductase deficiency. *Eur J Hum Genet* 1998; **6**: 257–265.
7 van der Put NMJ, Gabreëls F, Stevens EMB *et al*: A second common mutation in the methylenetetrahydrofolate reductase gene: An additional risk factor for neural tube defects? *Am J Hum Genet* 1998; **62**: 1044–1051.
8 Bailey LB, Gregory JF 3rd: Polymorphisms of methylenetetrahydrofolate reductase and other enzymes: metabolic significance, risks and impact on folate requirements. *J Nutr* 1999; **129**: 919–922.
9 Rosenblatt DS, Whitehead VM: Cobalamin and folate deficiency: acquired and hereditary disorders in children. *Semin Hematol* 1999; **36**: 19–34.
10 Chassé JF, Paul V, Escañez R, Kamoun P, London J: Human cystathionine β-synthase: gene organization and expression of different 5' alternative splicing. *Mamm Genome* 1997; **8**: 917–921.
11 Chen LH, Liu ML, Hwang HY, Chen LS, Korenberg J, Shane B: Human methionine synthase: cDNA cloning, gene localization, and expression. *J Biol Chem* 1997; **7**: 3628–3634.
12 Kraus JP, Oliveriusová J, Sokolová J *et al*: The human cystathionine β-synthase (CBS) gene: Complete sequence, alternative splicing, and polymorphisms. *Genomics* 1998; **52**: 312–324.
13 Leclerc D, Odievre M, Wu Q *et al*: Molecular cloning, expression and physical mapping of the human methionine synthase reductase gene. *Gene* 1999; **15**: 75–88.