Supplementary Materials

Transcription Start Site Identification

The following gene-specific primers were used in the 5′ RLM-RACE procedure:

MTHFD1 gene specific outer primer (Reverse) –

5'-GGCTTTGATCCCAATCTCTTCAGC-3'

MTHFD1 gene specific inner primer (Reverse) **–**

5'-CGCGAATTCATATTGCCAGGCGTGGTGTGAAAC-3'

Reporter Gene Constructs

The following list of PCR primers were used to generate the series of overlapping PCR products spanning 2kb upstream of the translational start site of the *MTHFD1* gene:

1.94kb Forward: CTCATGCCCATTATCCCAGCAT

1.94kb GatewayF: **GGGGACAAGTTTGTACAAAAAAGCAGGCTT**AAGGCCAAGGGAGACTA

1kb Forward: CCCAGGCAATTGTCCATCTAAC

1kb NheI: ATCGCTAGCCAACTGGCATGTA

0.59kb Forward: TTCCACATCTGCTGTCGAGTC

0.59kb SacI: CATGAGCTCTTCCACATCTGCTGTC

0.47kb NheI: ATCGCTAGCAAACCGGAGACTC

0.39kb SacI: TCGGAGCTCTGATTGGCTGGAATTAC

0.26kb SacI: ACGGAGCTCCATCACCGATTTTCTTC

0.11kb SacI: ATAGAGCTCATCCCCTGGCCAGTC

Promoter Reverse: TTATTAGTCCGCTGCCCACGA

BglII Reverse: CGCAGATCTTTAGTCCGCTGCC

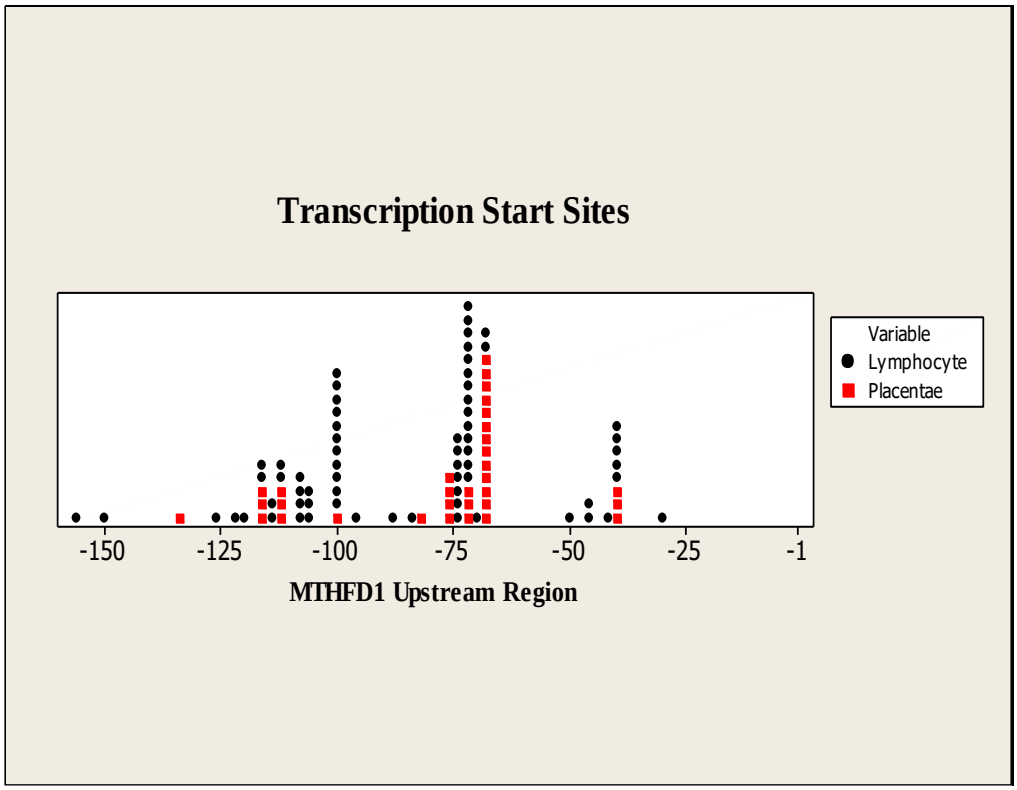
1.94kb GatewayR: **GGGGACCACTTTGTACAAGAAAGCTGGGT**CTTATTATTAGTCCGCTGC

Polymorphism Screening

A region encompassing 2kb upstream of the translational start site was screened for polymorphisms by a direct sequencing approach. The primer pairs used to amplify each overlapping section are detailed in the table below along with the most optimal annealing temperature.

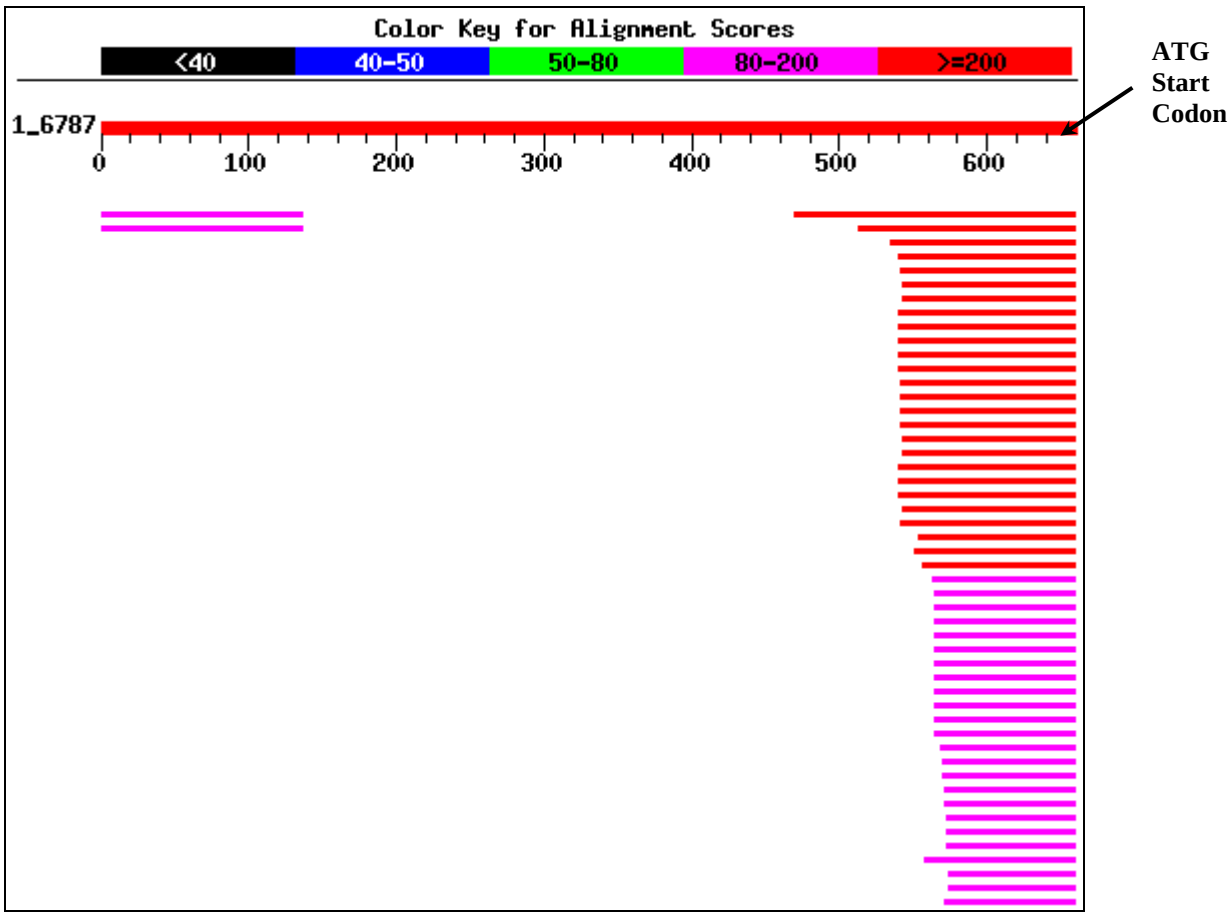
Primer Name	**Primer Sequence**	**Annealing (°C)**
Section 1 Forward	TTCCCGCCGAATACAAAGGC	57°C
Section 1 Reverse	TTATTAGTCCGCTGCCCACGA	
Section 2 Forward	TGGCGTAGGTGTGTGACAAA	58°C
Section 2 Reverse	CGGAAATGATTGGCGGAA	
Section 3 Forward	GGTCAACTGGCATGTAGCTTAC	54°C
Section 3 Reverse	ACTCGACAGCAGATGTGGAAC	
Section 4 Forward	CAGAATGAGACTCCGTCTCAAA	56°C
Section 4 Reverse	AACTCACGCGGTCTTCGTA	
Section 5 Forward	ATCAGTTTTAGGCCAGGTGC	56°C
Section 5 Reverse	CCGTTTACTACTGGCAGAGC	
Section 6 Forward	AGGCCAAGGGAGACTAATC	59°C
Section 6 Reverse	CCAGAGTGCTGGGATTACA	
Intron 1 Forward	CGCGTAAGCACCTGACATTGT	58°C
Intron 1 Reverse	GTCAAGCCTTGTCACCTCACTA	

Supplementary Figures



Supplementary Figure 1 *MTHFD1* transcription start site positions relative to the ATG start codon (+1)

Transcription start sites (TSSs) have been identified in at least 26 different positions over a 126 bp region in the *MTHFD1* upstream region from a total of five experiments using both lymphocyte and placental mRNA. Major start sites are seen at positions 68 bp, 72 bp, and 100 bp upstream of the ATG start codon. Different patterns of initiation are not evident between the different tissues investigated, nor were they evident between different three lymphocyte RNA samples representing three different individuals (not shown).



Supplementary Figure 2 BLAST alignment of 660bp of the *MTHFD1* upstream region with the NCBI human EST database (dbEST).

660 bp of the *MTHFD1* sequence upstream from the ATG start codon was aligned with clones from dbEST (Genbank release 147) to identify the clone with the most 5′ end that would be a likely transcription start site (TSS) or represent an alternate upstream exon. Most clones appear to start approximately 120 bp upstream (-120) of the ATG start codon. More detailed analyses revealed clones with a start site at almost every nucleotide from -120 to the start codon, indicating either a multiple TSS pattern or the existence of many clones lacking intact 5′ ends. The “purple” ESTs aligned on the extreme left above represent an exon from a predicted upstream gene, similar to the mouse Tex21 gene, rather than an alternate *MTHFD1* exon.