Alternative transcription initiation and splicing variants of the DHRS4 gene cluster

Qiaoxia ZHANG, Yifan LI, Gefei LIU, Xiaoyuan XU, Xuhong SONG, Bin LIANG, Rui LI, Jianping XIE, Mudan DU, Lifeng XIAO, Xueqiong GAN and Dongyang HUANG

Center for Molecular Biology, Shantou University Medical College, 22 Xinling Road, Shantou, Guangdong, 515041, China

Synopsis

The DHRS4 (short-chain dehydrogenase/reductase superfamily member 4) gene cluster, consisting of DHRS4 and its copy gene DHRS4L2, is localized on 14q11.2. The DHRS4 gene product NADP(H)-dependent retinol oxidoreductase participates in the metabolism of retinoids. The expression patterns of the DHRS4 gene cluster were investigated in human neuroblastoma cells. Transcript analysis of the DHRS4 gene cluster using 3'- and 5'-RACE (rapid amplification of cDNA ends), reverse transcription-PCR and bioinformatics approaches showed an alternative transcription start site in the copy gene DHRS4L2 which generates two transcripts, DHRS4A1 (GenBank® nucleotide sequence database accession number AY616183) and DHRS4A2 (AY943857), together with at least six alternative splicing variants (DHRS4A_v1–6) (AY920361, AY920362, DN237886, DN237887, DN237890 and DN237892 respectively), resulted from alternative splicing. DHRS4A1 and DHRS4A2 were specifically transcribed in neuroblastoma cells. RNA structural analysis of DHRS4A1 and DHRS4A2 suggested that they are non-coding RNAs. Expression analysis of DHRS4 by quantitative real-time PCR and Western blotting showed a lack of correlation between the levels of transcription and translation in the tissues examined. Bisulfite genomic sequencing PCR experiments indicated that the expression of DHRS4L2 was regulated by methylation of its CpG islands.

Key words: alternative splicing, DNA methylation, short-chain dehydrogenase/reductase superfamily member 4 (DHRS4) gene cluster, transcription start site (TSS)

INTRODUCTION

DHRS4 (short-chain dehydrogenase/reductase superfamily member 4) is an NADP(H)-dependent retinol oxidoreductase [1]. Because of its higher retinal reduction activity at physiological pH than other enzymes of the short-chain dehydrogenase/reductase superfamily [1,2], DHRS4 is considered to play an important role in maintaining the homeostasis of retinal and RA (retinoic acid). The expression of the DHRS4 gene cluster is regulated by alternative splicing: an alternative splicing variant (GenBank® nucleotide sequence database accession number AY071856) of the DHRS4 gene was cloned from human liver tissue [3,4]. Alternative splicing is reported to be associated with various disease processes, including tumour formation [5–7]. Differentially expressed splicing factors, mutations in cis-acting splicing elements and changes in the activity of regulatory proteins are thought to be involved in tumour development and progression, as they can compromise the accuracy of either constitutive or alternative splicing and sequentially deregulate crucial cellular processes, such as adhesion, proliferation, differentiation, death, motility and invasion [8,9].

Neuroblastoma is one of the most common solid malignant tumours occurring in children. Different signalling pathways are involved in the differentiation course of human neuroblastoma cells. RA can induce morphologic differentiation and growth inhibition of human neuroblastoma cells by inhibiting the proteasomal degradation of RA receptor α [10,11]. Endogenous (physiological) obstruction of RA generation may, therefore, be relevant to the abnormality of neural cells.

Considering the role of NADP(H)-dependent retinol oxidoreductase in the metabolism of retinoids and the effect of RA on human neuroblastoma cells, expression of the DHRS4 gene cluster in neuroblastoma cells deserves further study. Therefore, in the present study, the structure and transcription, and, in particular, the alternative splicing of the DHRS4 gene cluster in the human
neuroblastoma cell lines have been investigated. Our results not only uncovered the structure of the DHRS4 gene cluster, but also demonstrated the presence of novel alternative splicing variants with an alternative TSS (transcription start site).

MATERIALS AND METHODS

Cell lines and culture
Human neuroblastoma cell lines SK-N-SH, SK-SY-5Y and KP-N-NS were obtained from the Cell Bank of Chinese Academy of Sciences (Shanghai, China). The cell lines were maintained in RPMI 1640 medium supplemented with 10% (v/v) fetal bovine serum, 100 i.u./ml penicillin and 100 μg/ml streptomycin in a humidified 37°C incubator with 5% CO2. Cells which had undergone no more than six passages were used in all experiments.

Genomic DNA, total RNA and protein preparation
Genomic DNA, total RNA and total protein of human neuroblastoma cells were extracted from exponentially growing cells using the Genomic DNA Purification Kit (Promega), TRIzol® reagent (Invitrogen) and a mammalian protein extraction reagent (Pierce) respectively, following the manufacturer’s instructions.

Genomic DNA, total RNA and total protein of normal human tissues (brain, oesophagus, liver, placenta and testis) were purchased from BioChain.

3′-RACE (rapid amplification of cDNA ends)
3′-End amplification of alternative splicing variants was carried out using the 3′-RACE system (Invitrogen) according to the manufacturer’s protocol, with slight modifications in performing nested PCR.

The adaptor primer (Table 1) was used for the first-strand cDNA synthesis reaction. Target cDNA was amplified with the sense primer A-3p1 (for DHRS4A1 and DHRS4A2) and the antisense primer AUAP (abridged universal amplification primer) for the first PCR, whereas the sense primer A1-3p2 (for DHRS4A1) or A2-3p2 (for DHRS4A2) and the antisense primer AUAP were used for the second round of PCR (Table 1).

5′-RACE
RNA ligase-mediated 5′-RACE, which selects capped RNA, ensuring the amplification of full-length mRNA molecules, was performed using a Gene Racer RACE kit (Invitrogen). Truncated mRNA and non-mRNA species were eliminated from the RACE reaction by dephosphorylation of their 5′-phosphates with calf intestinal phosphatase. Capped mRNA was then decapped with tobacco acid phosphatase, exposing the 5′-phosphates required for oligonucleotide ligation. The ligated mRNA was reverse-transcribed with a designed gene-specific primer A-5p1 (Table 1). cDNAs were amplified by 5′-RACE PCR using the GeneRacer 5′ forward primer and a gene-specific reverse primer A-5p2 (Table 1). The PCR products were analysed by agarose-gel electrophoresis (1.5% gel).

Cloning and sequencing
PCR products of interest were purified using the Gel Extraction Mini Kit (Watson Biotechnologies) and cloned into the pGEM-T Easy vector (Promega). The sequence identity of each product was confirmed by DNA sequencing of three independent clones using the dideoxynucleotide method in an Applied Biosystems 3100 DNA sequencer.

Northern blot analysis
The Northern blot experiments were performed essentially following a standard protocol. Briefly, total cellular RNA (40 μg/lane) was denatured and electrophoresed in formaldehyde–agarose gels. The gels were evaluated by ethidium bromide staining, with the ribosomal 18 S and 28 S RNA bands used as molecular-mass markers. The separated RNA samples were transferred from the gels on to Hybond N nylon membranes (Amersham Biosciences) by electrophoretic transfer with TAE (Tris/acetate/EDTA) buffer [20 mM Tris/HCl (pH 7.8) 10 mM sodium acetate and 0.5 mM EDTA] and UV cross-linked to the membranes. The membranes were pre-hybridized for 1 h at 42°C in hybridization buffer [6 × SSC (1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), 5 × Denhardt’s solution (where 1 × Denhart’s solution is 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone and 0.02% BSA), 0.5% SDS, 100 μg/ml denatured salmon sperm DNA and 50% (v/v) formamide]. Hybridization was carried out at 42°C for 18–20 h with

### Table 1 3′-RACE and 5′-RACE primers used in the present study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5′–3′)</th>
<th>Gene</th>
<th>Use</th>
</tr>
</thead>
</table>
| Adaptor primer            | GGCACGCGTCAGCTAGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
amplified products were DNA sequenced as described above.

DHRS4L2 using the DHRS4
tions. PCR amplification of bisulfite-treated DNA was carried out (Chemicon International) following the manufacturer’s instruc-

Ea and Ea were obtained by RT-PCR (reverse transcription-PCR) using levels. cDNAs for Ea (a novel exon, see below) and produce control probes for basal general mRNA production (glyceraldehyde-3-phosphate dehydrogenase) was used to

membranes were washed twice in 2 × 0.1 SSC and 0.5% SDS at 60◦C for 5 min, followed by two washes in 0.1 × SSC and 0.1% SDS at 60◦C for 15 min.

cDNA probes, labelled by the random primer procedure using [α32P]dCTP (Multiprime labelling kit; Amersham Biosciences), and dissolved in hybridization buffer to yield an activity of approx. 2 × 106 c.p.m. of labelled probe/ml. cDNA for GAPDH (glyceralddehyde-3-phosphate dehydrogenase) was used to produce control probes for basal general mRNA production levels. cDNAs for Ea (a novel exon, see below) and GAPDH were obtained by RT-PCR (reverse transcription-PCR) using Ea and GAPDH primers (Table 2). After hybridization, the membranes were washed twice in 2 × SSC and 0.5% SDS at room temperature (25◦C) for 5 min, followed by two washes in 0.1 × SSC and 0.1% SDS at 60◦C for 15 min.

qRT-PCR (quantitative real-time PCR) analysis
Total RNA was reverse-transcribed into cDNA using the SuperScript® III First-Strand Synthesis System (Invitrogen). PCR was performed in triplicate using the SYBR Green PCR Master Mix (Invitrogen) with cDNA (3 μl) and DHRS4, Ea, DHRS4L2 and GAPDH primers (300 nM each) (Table 2) in a 50 μl (final volume) reaction mixture on an ABI 7000 Real-Time PCR System (Applied Biosystems). An RT-minus control was included to detect any contaminating genomic DNA. To exclude the occurrence of primer dimers and non-specific PCR products, melting-curve analysis was performed using the ABI PRISM 7000 Sequence Detection System Software version 1.0.1 (Applied Biosystems).

Bisulfite genomic sequencing PCR
Bisulfite treatment of DNA, which converts unmethylated cyto-
sine into uracil residues without affecting methylated cytosine residues, was done using the CpGenome™ DNA modification kit (Chemicon International) following the manufacturer’s instructions. PCR amplification of bisulfite-treated DNA was carried out using the DHRS4, Ea and DHRS4L2 primers (Table 3) and the amplified products were DNA sequenced as described above.

Production of a polyclonal anti-DHRS4 antibody
A rabbit polyclonal antibody against 30 kDa DHRS4 was raised as described previously [12]. Briefly, the expression vector pDEST17-DHRS4, containing sequence-confirmed full-length DHRS4, was constructed and transformed into Escherichia coli BL21-AI cells (Invitrogen). To overexpress DHRS4, 0.2% L-arabinose was added to a mid-exponential culture of BL21-AI [D600 (attenuance at 600 nm) of 0.6], and the culture was grown continuously for 3 h at 37◦C. The isolated overexpressed proteins were separated on a preparative SDS-polyacrylamide gel (12% gel) and visualized by immersing the gel in 0.25 M potassium chloride. The gel slice containing the desired protein band was mixed with Freund’s complete adjuvant and used for the immunization of rabbits. Rabbit anti-DHRS4 antisera was isolated by using a Protein G column on the ATKA Purifier 100 system (Amersham Biosciences) following the manufacturer’s guidelines. The sensitivity and specificity of the purified antibody was determined by dot blot and Western blot analysis.

Western blot analysis
Total protein of normal human tissues (brain, oesophagus, liver, placenta and testis) and SK-N-SH cells was resolved by SDS/PAGE (15% gels) and transferred on to a PVDF membrane (Immobilon P, 0.22-μm-pore-size; Millipore). The primary antibody was the rabbit polyclonal anti-DHRS4 antibody, and the secondary antibody was a goat anti-rabbit IgG conjugated to horseradish peroxidase (1:3000 dilution, Pierce). Immunoblots were visualized on photographic films using the SuperSignal West Pico chemiluminescent substrate (Pierce) following the manufacturer’s instructions.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’→3’)</th>
<th>Gene</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Fwd</td>
<td>CTGCTCTGATCCATGCACAAGGCGG</td>
<td>DHRS4</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>D-Rev</td>
<td>GCTGCTGACGACCACTGGGCCCC</td>
<td>DHRS4</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>A-Fwd</td>
<td>AGCAGACCGCCACAGTGCCT</td>
<td>Ea</td>
<td>qRT-PCR, Northern blot</td>
</tr>
<tr>
<td>A-Rev</td>
<td>GACGAGCTCCGTACCTACAACTC</td>
<td>Ea</td>
<td>qRT-PCR, Northern blot</td>
</tr>
<tr>
<td>L2-Fwd</td>
<td>AGGCCTCTGTGCTGGGACAGGAAG</td>
<td>DHRS4L2</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>L2-Rev</td>
<td>TGCTGCTTCCGGCTGCTGACACCA</td>
<td>DHRS4L2</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>G-Fwd</td>
<td>GAGGTGAAGGCTCGGAGTC</td>
<td>GAPDH</td>
<td>qRT-PCR, Northern blot</td>
</tr>
<tr>
<td>G-Rev</td>
<td>GAAGATGGTGATGGGATTTC</td>
<td>GAPDH</td>
<td>qRT-PCR, Northern blot</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’→3’)</th>
<th>Gene</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-Fwd</td>
<td>CGAGTTTTATAGGTAATTGAAAC</td>
<td>DHRS4, DHRS4L2</td>
<td></td>
</tr>
<tr>
<td>NM-Fwd</td>
<td>TGGATTTTATAGGTAATTGAAT</td>
<td>DHRS4, DHRS4L2</td>
<td></td>
</tr>
<tr>
<td>D-Rev</td>
<td>AAAACCTCGACGGAAC</td>
<td>DHRS4</td>
<td></td>
</tr>
<tr>
<td>D-NM-Rev</td>
<td>CAAAACATCCAACAAAAAC</td>
<td>DHRS4</td>
<td></td>
</tr>
<tr>
<td>A-M-Fwd</td>
<td>GGTGTATAAATAGGAATGGG</td>
<td>Ea</td>
<td></td>
</tr>
<tr>
<td>A-M-Rev</td>
<td>TAAACATCAAATAGCAACAA</td>
<td>Ea</td>
<td></td>
</tr>
<tr>
<td>L2-M-Rev</td>
<td>AAACCCTAAAAACTCGAC</td>
<td>DHRS4L2</td>
<td></td>
</tr>
<tr>
<td>L2-NM-Rev</td>
<td>AAACCCTAAAAACTCGAC</td>
<td>DHRS4L2</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Primers used in Northern blot and qRT-PCR analysis
Fwd, forward; Rev, reverse.

Table 3: Primers used in bisulfite genomic sequencing PCR
Fwd, forward; Rev, reverse.
RESULTS AND DISCUSSION

Structure of the human DHRS4 gene cluster
Examination of the DHRS4 gene cluster localized on chromosome 14q11.2 revealed that there are two gene copies, DHRS4 and DHRS4L2, which share 97.5% identity, with a 19540 nt fragment (chromosome 14: 23508327–23527866) inserted between them. The RefSeq of DHRS4 and DHRS4L2 recorded in the GenBank® nucleotide sequence database (accession numbers NM_021004 and NM_1980083) both consist of eight exons and all of the splicing junctions conform to the GT/AG rule (Figure 1a).

Identification of alternative splicing variants DHRS4A1 and DHRS4A2
RT-PCR and RACE were carried out to isolate alternative splicing variants of DHRS4L2 in SK-N-SH cells. Two cDNAs were obtained, a larger one of 1244 bp [DHRS4A1 (GenBank® nucleotide sequence database accession number AY616183)], and a smaller one of 1146 bp [DHRS4A2 (AY943857)]. A novel exon, termed Ea, which is approx. 19 kb upstream of the first exon (E1) of DHRS4L2 in genomic DNA, was identified by comparing sequences of these two cDNAs and the RefSeq of DHRS4L2. The presence of Ea would lead to the formation of a new intron of 19849 bp in length (chromosome 14: 23509382–23529230). DHRS4L2 exons E1, E4, E5 and E6 are absent in DHRS4A1 and an additional exon (E3) is absent in DHRS4A2 (Figure 1b). A similar splicing pattern was observed in a splicing variant of DHRS4 named NRDRiso (AY071856), where E4, E5 and E6 are absent [4].

Figure 2 Northern blotting analysis were probed using 32P-labelled Ea cDNA
The transcripts including Ea (DHRS4A1 or DHRS4A2) were detectable in all of the cell lines tested. A probe against GAPDH was used to monitor comparable loading between the samples. 18 S and 28 S ribosomal RNA bands are also indicated.

The expression of Ea was analysed in other two neuroblastoma cell lines, SK-SY-5Y and KP-N-NS, by Northern blot analysis. The results showed that the transcripts that contain Ea were also expressed in SK-SY-5Y and KP-N-NS cells, similarly to SK-N-SH cells (Figure 2).
Transcription of the DHRS4 gene cluster

Figure 3 Nucleotide sequence and structure of the 5′-flanking region of DHRS4L2

Exon sequences are represented in upper case, with intron sequences in lower case. The first and second exons of DHRS4A1 (Ea and E2) are shaded in black and the first exon of DHRS4L2 (E1) is shaded in grey. The localization of introns was deduced from the comparison of cDNA sequences (GenBank® nucleotide sequence database accession numbers AY616183 and AY943857) and the genomic sequences obtained from the human genome database (http://www.ncbi.nlm.nih.gov/Genbank/). Parts of the intron sequences are not presented; their respective positions in the genomic sequence and length are indicated by the number of bp in parentheses. CpG regions are double underlined and CpG dinucleotides are highlighted in bold font. Arrows indicate the positions of various PCR primers. Position of 5′-RACE products (angled double-line arrow) and computationally predicted TSSs by Eponine (angled single-line arrow) and FirstEF and NNPP2.2 (angled broken-line arrow) are shown.

RT-PCR was performed with total RNA from SK-N-SH cells in order to identify new splicing variants that contain Ea. Multiple RT-PCR fragments were detected with the sense primer specific for Ea and the antisense primers specific for E1–E8. By sequencing over 40 clones, we identified at least six alternative splicing variants, which were designated DHRS4A_v1–6.
Figure 4  Expression of the DHRS4 gene cluster at the transcript level as determined by qRT-PCR
Results are means ± S.E.M. (n = 3). B, brain; E, oesophagus; L, liver; P, placenta; SK, SK-N-SH cell line; and T, testis.

Table 4  DHRS4A1- and DHRS4A2-specific primers used in RT-PCR
Fwd, forward; Rev, reverse.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5′→3′)</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1Fwd</td>
<td>GCCTCGCGAGGAGGGTCACGGTGG</td>
<td>DHRS4A1</td>
</tr>
<tr>
<td>A1Rev</td>
<td>GTCCATCCAGGCTCTGCCCACACC</td>
<td>DHRS4A1</td>
</tr>
<tr>
<td>A2Fwd</td>
<td>TGAAGCGGAGTGCTGATGCTGGA</td>
<td>DHRS4A2</td>
</tr>
<tr>
<td>A2Rev</td>
<td>TCCTTGGCGACGAGCATGGCCA</td>
<td>DHRS4A2</td>
</tr>
</tbody>
</table>

Identification of TSS in DHRS4A1 and DHRS4A2
Two independent methods, computational analysis and RACE, were used to identify the TSS of the DHRS4A variants.

Analysis of 1.5 kb gene segment upstream of Ea by computational analysis using NNPP2.2 (Neural Network Promoter Prediction 2.2) [12a], Eponine [13] and FirstEF [14] software, which utilize various operating principles, including the presence of CpG islands, TATA, CCAAT and Inr (initiator) sequences, predicted that the transcription of Ea could start from two different TSSs. One TSS, predicted by NNPP2.2 (confidence interval ≥0.8) and FirstEF (confidence interval 1.0), was 119 nt upstream of Ea, whereas the other TSS, predicted by Eponine, was 92 nt upstream of Ea, with a confidence interval of ≥0.999.

Figure 5  DHRS4A1 (a) and DHRS4A2 (b) were specifically expressed in neuroblastoma cell lines as determined by RT-PCR
GAPDH is shown as a control. DNA ladder markers are shown on the right-hand side (M). B, brain; E, oesophagus; KP, KP-N-NS cell line; L, liver; P, placenta; SK, SK-N-SH cell line; SY, SK-SY-5Y cell line; and T, testis.
Transcription of the \textit{DHRS4} gene cluster

(a) Cell lysates from BL-21AI cells transformed with the pDEST 17 vector only (lane 1), the \textit{DHRS4A1} variants (lane 2) and the \textit{DHRS4A2} variants (lane 3), were separated by SDS/PAGE (15\% gels) and Western blotted using an anti-DHRS4 antibody, which recognized proteins translated from \textit{DHRS4A1} and \textit{DHRS4A2}. Molecular masses are indicated on the left-hand side [in kD (kDa)]. (b) Expression of the \textit{DHRS4} gene cluster at the protein level was detected by Western blotting analysis in normal human tissues and in SK-N-SH cells. Molecular masses are indicated on the left-hand side [in kD (kDa)]. B, brain; E, oesophagus; L, liver; P, placenta; SK, SK-N-SH cell line; and T, testis.

To confirm the presence of the predicted TSSs in Ea, we performed RNA ligase-mediated 5'-RACE, which ensures the amplification of full-length mRNA molecules. Cloning and subsequent sequencing of the PCR amplification product identified the first nucleotide of Ea (chromosome 14: 23508988) as a TSS (Figure 3).

The location of the TSS of \textit{DHRS4A1} and \textit{DHRS4A2}, which is approx. 19 kb upstream of the TSS of the \textit{DHRS4L2} RefSeq, indicates the presence of an alternative promoter for \textit{DHRS4L2}. Alternative promoter selection is a well-recognized mode of regulation of mammalian gene expression [15]. In addition to alternative splicing, it is a strategy used for increasing the functional diversification of human genes [16,17]. With alternative promoters that consist of different modules of transcriptional regulatory elements, diversified transcriptional regulation should be enabled at a single locus. The combinatory use of alternative splicing and alternative promoters would increase even further the potential complexity of the products expressed from a single gene. Therefore the new TSS would provide the \textit{DHRS4L2} gene with the possibility of both multifunction or multiregulatory modes.

**Expression level of the \textit{DHRS4} gene cluster**

Ea expression is not exclusive to human neuroblastoma cell lines; its expression at a transcriptional level has been found in oesophageal tumour tissue (GenBank® EST sequence database accession number DB016984) as well as some normal human tissues, such as brain (DA776573), hypothalamus (BC041992, BL602725, DB489612 and DB499846), retina (BX647440), cerebellum (DA061260), placenta (BX380415) and testis (BG723975, BG724049 and CD244539).

The transcriptional level of \textit{DHRS4}, Ea and \textit{DHRS4L2} in normal human tissues (brain, oesophagus, liver, placenta and testis) and SK-N-SH cells was measured by qRT-PCR, with the \textit{GAPDH} gene used as an internal control. As shown in Figure 4, Ea and \textit{DHRS4} were transcribed abundantly in testis and liver respectively. The primers for \textit{DHRS4L2} used in the present study had been tested by amplifying genomic DNA and no amplification of \textit{DHRS4L2} [\(C_T\) (threshold cycle value) = 40] was observed.

Although Ea was also transcribed in these normal human tissues, the transcripts \textit{DHRS4A1} and \textit{DHRS4A2} were expressed specifically in human neuroblastoma cell lines. The transcriptional levels of \textit{DHRS4A1} and \textit{DHRS4A2} were measured by RT-PCR using \textit{DHRS4A1}- and \textit{DHRS4A2}-specific primers A1Fwd/A1Rev and A2Fwd/A2Rev (Table 4), with the \textit{GAPDH} gene used as an internal control. As shown in Figure 5, \textit{DHRS4A1} and \textit{DHRS4A2} were transcribed specifically in human neuroblastoma cell lines, but neither of them was detected in normal human tissues (brain, oesophagus, liver, placenta and testis).
The variants *DHRS4A1* and *DHRS4A2* have no functional translation start site in the RefSeq of *DHRS4L2* as they lack E1. Their deduced translational products of 91 and 57 amino acid residues respectively are available in the NCBI Entrez Protein database (accession numbers AAT70758 and AAX49568). Their amino-acid sequences were highly similar (95% and 98% identity respectively) to the corresponding amino-acid sequence of the *DHRS4* protein.

For detection of the deduced proteins, the anti-DHRS4 antibody was tested for its ability to recognize the shorter proteins encoded by *DHRS4A1* and *DHRS4A2*. The expression vectors pDEST17-A1 and pDEST17-A2, containing *DHRS4A1* and *DHRS4A2* coding sequences respectively, were constructed and transformed into *E. coli* BL21-AI cells and the cell lysate was used for Western blot analysis. The result showed that the antibody was functioning satisfactorily (Figure 6a). DHRS4 protein was detected in liver and testis by Western blot analysis using the rabbit polyclonal anti-DHRS4 antibody (Figure 6b). However, the deduced proteins encoded by *DHRS4A1* and *DHRS4A2* (with theoretical molecular masses of 9.8 kDa and 6.3 kDa respectively) were not observed in SK-N-SH cells or in the normal tissues examined (Figure 6b).

The absence of the translational products from *DHRS4A1* and *DHRS4A2* splicing variants led us to the analysis of their secondary structure. Analysis using RNAstructure4.5 [18] showed that the *DHRS4A1* and *DHRS4A2* isoforms can form a stable double-stranded RNA structure. These observations indicate that the alternative splicing variants are ncRNAs (non-coding RNAs) acting as RNA molecules. The finding of ncRNAs has been one of the most startling discoveries in the genomic era. Except for siRNAs (small interfering RNAs) and miRNAs (microRNAs), ncRNAs are very long and highly abundant, but are the least-characterized type of RNA, and little is known about their function [19]. As seen in Figure 4 and Figure 6(b), the levels of RNA transcript do not correlate well with the levels of expressed proteins in many tissues examined. For instance, despite having higher expression...
of RNA than the testis, there was no detectable level of DHRS4 protein in the oesophagus and placenta. The observed discrepancies between transcription and translation suggest a potential regulatory role of the products of the DHRS4 gene cluster in the differential expression of DHRS4 proteins in the various tissues examined.

CpG islands and their methylation status in the DHRS4 gene cluster

We identified one CpG island with 43 CpGs distributed throughout the region in the potential promoter region of DHRS4A1 and DHRS4A2, and two other islands around the first exon of DHRS4 (e1) with 32 CpGs and E1 of DHRS4L2 with 18 CpGs (Figure 7). Cytosine methylation patterns of CpG dinucleotides are heritable, but this genetic modification is potentially reversible. Cytosine methylation within the context of CpG dinucleotides is a conserved epigenetic silencing mechanism involved in the regulation of gene expression [20,21]. Acquired epigenetic abnormalities are considered to be related to cancer [22]. Aberrant hypermethylation of CpG islands in the promoter region in many cancer-related genes results in the silencing of their expression [23,24].

The methylation status in the regions covering the area of greatest CpG density, from nt –194 to nt 175 of e1, nt –239 to nt 177 of Ea and nt –110 to nt 74 of E1, were analysed by bisulfite genomic sequencing PCR. Almost complete methylation was seen in the CpG island of DHRS4L2, whereas the CpG islands of Ea and DHRS4 were almost unmethylated (Figure 7).

The expression level of the DHRS4 gene cluster and the methylation status of its CpG islands strongly suggested that the transcription of the RefSeq of DHRS4L2 is controlled by cytosine methylation of the CpG island, whereas regulatory mechanisms other than cytosine methylation may be involved in Ea and DHRS4 expression. Takai et al. [25] have reported that although 60% of human genes have CpG islands in the promoter region or the first exon, more than 80% of all CpG islands are not related to genes and thus are unlikely to regulate gene expression.

In conclusion, we have identified novel alternative splicing variants which are transcribed from an alternative TSS within the DHRS4 gene cluster. Although it is possible that the progression of human neuroblastoma leads to disturbances in the splicing machinery and, in return, produce alternative splicing variants, the neuroblastoma-specific variants DHRS4A1 and DHRS4A2 could be useful markers for neuroblastoma.

ACKNOWLEDGEMENTS

We thank William Ba-Thein for comments on the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China [grant number 30700474]; and the Natural Science Foundation of Guangdong Province [grant number 8251503102000003].

REFERENCES