Mouse cytosolic and mitochondrial deoxyribonucleotidases: cDNA cloning of the mitochondrial enzyme, gene structures, chromosomal mapping and comparison with the human orthologs[☆]

Chiara Rampazzo[a], Maria Kost-Alimova[b], Benedetta Ruzzenente[a],
Jan P. Dumanski[c], Vera Bianchi[a],*

[a]*Department of Biology, University of Padua, Via Ugo Bassi 58B, I 35131 Padua, Italy*
[b]*Microbiology and Tumor Biology Center (MTC), Karolinska Institute, S 17177 Stockholm, Sweden*
[c]*Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, S 75185 Uppsala, Sweden*

Abstract

Two of the five known mammalian $5'$-nucleotidases show a preference for the dephosphorylation of deoxynucleoside-$5'$-phosphates. One is a cytoplasmic enzyme (dNT-1), the other occurs in mitochondria (dNT-2). The human mitochondrial enzyme, recently discovered and cloned by us, is encoded by a nuclear gene located on chromosome 17 p11.2 in the critical region deleted in the Smith–Magenis syndrome (SMS), a genetic disease of unknown etiology. Looking for a model system to study the possible involvement of dNT-2 in the disease, we have cloned the cDNA of the mouse ortholog. The deduced protein sequence is 84% identical to the human ortholog, has a very basic NH_2-terminus, a very high calculated probability of being imported into mitochondria and contains the DXDXT/V motif conserved among nucleotidases. Expression in *Escherichia coli* of the predicted processed form of the protein produced an active deoxyribonucleotidase. We also identified in genomic sequences present in the data base the structures of the murine genes for the cytosolic and mitochondrial deoxyribonucleotidases (*Nt5c* and *Nt5m*). PAC clones for the two loci were isolated from a library and used for chromosomal localization by fluorescent in situ hybridization. Both genes map on chromosome 11: *Nt5c* at 11E and *Nt5m* at 11B, demonstrating the presence of the dNT-2 locus in the mouse *shaker-2* critical region, the murine counterpart of the human SMS region. We performed pair-wise dot-plot and PIP (percent identity plot) analyses of mouse and human deoxyribonucleotidase genes, and found a strong conservation that extends also to some intronic sequences of possible regulatory significance. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Deoxyribonucleotide pool; $5'$-Nucleotidase; Smith–Magenis syndrome; Mouse *shaker-2* region; Fluorescent in situ hybridization; Synteny

1. Introduction

$5'$-Nucleotidases catalyse the cleavage of the phosphoester bond of nucleoside $5'$-monophosphates producing nucleosides and inorganic phosphate. In mammalian cells they form a heterogeneous group of enzymes differing in tissue specificity, subcellular location, primary structures and substrate specificity. Originally they were discovered from their enzyme activity, more recently via cloning of their cDNAs. One enzyme is an ubiquitous ectonucleotidase anchored to the surface of the plasma membrane (Resta et al., 1998). Four enzymes are localized in the cytosol. Two of them occur ubiquitously but differ in substrate specificity, the 'high K_m cytosolic nucleotidase' (cN-II) (Itoh, 1993) and the $5'(3')$-deoxyribonucleotidase (dNT-1) (Rampazzo et al., 2000a). The remaining two cytosolic enzymes, cN-I (Sala-

Newby et al., 1999) and P5′N-I (Paglia et al., 1984), are largely tissue-specific. Finally, we recently discovered a 5′-nucleotidase in mitochondria (dNT-2) with a high specificity for dUMP and dTMP (Rampazzo et al., 2000b).

A common feature of 5′-nucleotidases is their rather broad substrate specificity which in many cases overlaps between different enzymes. Their physiological function is understood incompletely, although their enzyme activity suggests that they are involved in the turnover of nucleotide pools and hence in the regulation of the many nucleotide-dependent reactions. By now all the above enzymes have been cloned, and thus it is possible to manipulate their activity in situ and to investigate directly their role in cell metabolism.

We are interested in the regulation of the size of deoxyribonucleoside triphosphates (dNTP) pools in relation to DNA synthesis. The allosteric regulation of ribonucleotide reductase – the enzyme responsible for the de novo synthesis of deoxyribonucleotides – is of major importance in this respect (Reichard, 1988). We have proposed that, in addition, specific 5′-nucleotidases also are involved by participating in substrate or futile cycles together with deoxynucleoside kinases (Bianchi et al., 1986). From results of isotope flow experiments we suggested that dNT-1 together with the cytosolic thymidine kinase (TK1) participates in such a cycle regulating the intracellular concentration of deoxyuridine and thymidine nucleotides (Gazziola et al., 2001).

The finding of a mitochondrial deoxyribonucleotidase gives an occasion to investigate the regulation of intramitochondrial dNTP pools. Mitochondria replicate their DNA independently of nuclear DNA synthesis in both cycling and differentiated cells (Shadel and Clayton, 1997) and must control their dNTP pools although the surrounding cytosol is a variable source of precursors. We propose that dNT-2 together with the mitochondrial thymidine kinase (TK2) (Johansson and Karlsson, 1997) creates an intramitochondrial substrate cycle involved in the regulation of the mitochondrial dTTP pool. The enzyme is also of interest because its gene is located on human chromosome 17 in the minimal deletion region defining the hereditary Smith–Magenis syndrome (Chen et al., 1996). This critical region contains at least eight genes and it is not known if and how they contribute to the pathological phenotype. Besides the manipulation of enzyme activity in cultured cells, the construction of mouse models is at present the most informative approach to investigate the physiological function of a new gene. The latter approach requires knowledge of the genomic structure of the gene. The present work provides this knowledge and compares the structure and chromosomal localization of the genes encoding the cytosolic dNT-1 and the mitochondrial dNT-2 in humans and mice. Evolutionary analyses of the loci revealed strong conservation of the structure of the four genes and islands of conservation in non-coding regions that may be regulatory elements.

2. Materials and methods

2.1. Identification of the murine dNT-2 cDNA, cloning and expression in bacteria

We screened a cDNA library derived from *Mus musculus* diaphragm (Lambda cDNA library in Zap XR Vector; Stratagene, Heidelberg, Germany) with MNT5 and MNT4, two gene-specific primers (see Table 1 for all the primers used in this study). Two overlapping fragments of the cDNA were amplified by PCR by using MNT4 in combination with the vector-specific primer A and primer MNT5 with the vector-specific primer B (Stratagene). The cDNA was cloned into pGEM-T Easy Vector (Promega, Madison, WI, USA) and sequenced on both strands. Using primers MNT11 and MNT12 that introduced a 5′ *Nde*I and a 3′ *Bam*HI site, respectively, we amplified a truncated cDNA lacking the first 21 codons but with an additional methionine codon in 5′ and subcloned it into pET20-b (Novagen, Madison, WI, USA) to give plasmid p2M-mdNT-2. We used this plasmid to express in bacteria the recombinant protein lacking the first 21 amino acids that include the putative mitochondrial targeting signal (Neupert, 1997). The construct was transformed into *Escherichia coli* BL21(DE3) p*lysS* and the bacterial cultures grown at 37 °C were induced for 3 h with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG). In extracts prepared as described (Rampazzo et al., 2000b) we detected the recombinant protein by 12% SDS–PAGE and by an enzymatic assay with [^{3}H]dUMP as the substrate (Rampazzo et al., 2000b). Assays were done with at least two different aliquots of bacterial extracts to assure proportionality. Enzyme specific activity is units/mg protein (1 unit = 1 μmol of deoxyuridine formed per minute).

2.2. Isolation of murine genomic clones including the Nt5c and Nt5m genes

High-density colony hybridization filters from RPCI 21 female mouse PAC library (segment 2, composed of 132473 individual clones, 147 kb average insert size from strain 129S6/SvEvTac) (http://www.chori.org/bacpac/) (Osoegawa et al., 2000) were screened with probes for *Nt5c and Nt5m*. To screen the mouse PAC libraries we used a frag-

Table 1
Sequences of PCR primers used in this study

Name	Primer sequence (5′–3′)
CR4	GCGGATAAAGTGGCCAGTGTGTACGAAGC
CR2	TGTATTTCCATGCCAGCCAGGGTCCAGG
MNT1	CCTGGCAGGCATGTA
MNT2	ATTTTTCATATTGGCAACA
MNT4	AGTAGAGGTACCTACATGCCTGCC
MNT5	GAGGACTGTACTATGGCTCTCTCGTG
MNT11	CATACATATGCGGGCGAGCAGCCGCGC
MNT12	CTGAGGATCCGAGAGAGCCATAGTACAG
A	TCCGGCTCGTATGTTGTGTGGAAT
B	GTTGTAAAACGACGGCCAGTGAAT

ment of the human dNT-1 cDNA derived from IMAGE cDNA clone 526727 amplified with primers CR4 and CR2 (Table 1) and a fragment from the mouse dNT-2 cDNA derived from the amplification of the IMAGE cDNA clone 330411 with primers MNT1 and MNT2 (Table 1). The cDNA probes were tested on human and murine Southern and Northern blots, prior to their use for the screening of murine PAC libraries. Probes were labeled by the random priming method (Feinberg and Vogelstein, 1984), filters were hybridized, stringently washed and exposed to X-ray films according to standard protocols (Sambrook et al., 1989). The positive clones were selected from the plates of the RCPI21 library and grown separately on agar plates with kanamycin. Three independent bacterial colonies were picked for each positive clone and PAC DNA was prepared using the Qiagen method from each colony, followed by *Eco*RI restriction endonuclease digestion and agarose gel electrophoresis (Sambrook et al., 1989).

2.3. Chromosomal mapping of Nt5c and Nt5m by fluorescent in situ hybridization

Clones of murine genomic DNA were used as probes for fluorescent in situ hybridization. For chromosomal localization of the genes we used clone 476P6 (*Nt5m* mapping) and clone 432L13 (*Nt5c* mapping). The two clones were used in two-color FISH on mouse metaphase chromosomes prepared from BALB/c embryonic fibroblasts as described previously (Fedorova et al., 1997). We labeled DNA isolated from clone 476P6 with biotin-14-dATP using a BioNick Labeling System (Gibco BRL, Bethesda, MD, USA.) and developed the hybridization signal using Fluor-Link Cy3 labeled streptavidin (Amersham Pharmacia Biotech, UK). To label DNA isolated from clone 432L13 with DIG-11-dUTP we used a DIG-Nick Translation Mix (Roche Molecular Biochemicals, Mannheim, Germany) and developed the hybridization signal using Anti-digoxigenin-fluorescein (Roche Molecular Biochemicals). The chromosome carrying FISH signals was identified in the same metaphases by a second hybridization with mouse chromosome 11 specific painting probe (Cambio, Cambridge, UK). Analysis of FISH results was done with a fluorescence microscope (Leitz-DMRB, Leica, Heidelberg, Germany) equipped with a Hamamatsu 4800 cooled CCD camera (Hamamatsu, Herrsching, Germany) and the computer program for image analysis Adobe Photoshop 5.5 (Adobe Systems, San Jose, CA, USA).

2.4. Bioinformatics

We used the Blast family of programs for data base searches on the NCBI/NIH server (www.ncbi.nlm.nih.gov/BLAST) to identify mouse EST clones homologous to *NT5M*. EST clones from Research Genetics Inc. (Huntsville, AL, USA) were resequenced by us before use. The genomic structures of *NT5C*, *Nt5c* and *Nt5m* were identified in clones present in GenBank with the accession numbers AC022211, AL645470 and AC068808, respectively, and deposited in the Third Party Annotation Section of the DDBJ/EMBL/GenBank data bases (accession numbers BK000192, BK000208 and BK000190).

Clone AL645470 appeared in the GenBank data base only when our work was almost completed. Therefore, for our analyses of the genomic organization of the human and mouse deoxynucleotidase genes we referred to the Celera data bases. Celera's data base of human genomic sequences was searched with human sequences containing *NT5C* and *NT5M* (accession numbers AC022211 and AC006071, respectively). Celera's data base of mouse genomic sequences was searched with clone AC068808, containing *Nt5m*. The *Nt5c* sequence was found with the BlastN program on Celera's server using the mouse cDNA sequence (accession number AF078840). The sequences which were identified after searching the Celera server were submitted to the PIP-maker server (http://nog.cse.psu.edu/pipmaker/) (Schwartz et al., 2000).

3. Results and discussion

3.1. Cloning of the murine dNT-2 cDNA

A search of the mouse EST data base for sequences homologous to the human dNT-2 cDNA recently cloned in our laboratory (Rampazzo et al., 2000b) produced several positive hits, but all the ESTs found lacked the $5'$-end of the coding sequence. On the basis of the available information we designed two specific primers (MNT4 and MNT5, Table 1) and used them to screen a mouse cDNA library. The screen lead to the isolation of a 1290 nt long cDNA containing the coding sequence for the mouse dNT-2. The deduced protein sequence is 220 amino acids long (Fig. 1) and has a very basic NH$_2$-terminus (isoelectric point 11.70). MitoProt II analysis (http://www.mips.biochem.mpg.de/proj/medgen/mitop/) (Claros and Vincens, 1996) predicts mitochondrial import of the protein with very high probability (0.9738). Alignment of the amino acid sequence with those of the human dNT-2 and the human and mouse dNT-1 (Fig. 1) shows a high similarity among the four proteins ($\cong$ 50% identity). All contain in their NH$_2$-terminus the DMDGV sequence, a conserved DXDX(T/V) motif recently indicated as a nucleotidase signature and shown to be directly involved in phosphotransfer during catalysis (Allegrini et al., 2001). It is noteworthy that a truncated form of mouse dNT-1 starting from methionine 13, i.e. the methionine present in the above conserved sequence, lacked enzyme activity (Rampazzo et al., 2000a). The mitochondrial targeting signals of the human and mouse dNT-2s are underlined in Fig. 1 where a bar indicates the putative cleavage site at which the pre-protein is processed during import (Neupert, 1997; Branda and Isaya, 1995). The underlined mouse signal diverges from the 'typical' mitochondrial targeting sequence at position -5 where R is present instead of T/S/G and at position -8 where an A substitutes the expected F/L/I. The latter substitution occurs

```
hdNT-2    1  MIRLGGWCARRLCSAAVPAGRRGAAGGLGLA GGRALRVLVDMDGVLADFEGGFLRKFRAR
mdNT-2    1  -------MHRLRGCCARPRGAPLRAERS RASSRALRVLVDMDGVLADFEGGFLRKFRAR
hdNT-1    1  -----------------------------MARSVRVLVDMDGVLADFEAGLLRGFRRR
mdNT-1    1  -----------------------------MAVKRPVRVLVDMDGVLADFESGLLQGFRRR

hdNT-2   61  FPDQPFIALEDRRGFWVSEQYGRLRPGLSEKAISIWESKNFFFELEPLPGAVEAVKEMAS
mdNT-2   53  FPDLPFVALEDRRGFWVSEQYGRLQPGLSEKAISIWESKDFFFELEPLPGAVEAVKQMAN
hdNT-1   30  FPEEPHVPLEQRRGFLAREQYRALRPDLADKVASVYEAPGFFLDLEPIPGALDAVREMND
mdNT-1   32  FPEEPHVPLEQRRGFLANEQYGALRPDLAEKVASVYESPGFFLNLEPIPGALDALREMND
                                           ▲
hdNT-2  121  LQNTDVFICTSPIKMFKYCPYEKYAWVEKYFGPDFLEQIVLTRDKTVVSADLLIDDRPDI
mdNT-2  113  LQNTDVFICTSPIKMFKYCPYEKYAWVEKHFGPDFLEQIVLTRDKTVISADLLIDDRPDI
hdNT-1   90  LPDTQVFICTSPLLKYHHCVGEKYRWVEQHLGPQFVERIILTRDKTVVLGDLLIDDKDTV
mdNT-1   92  MKDTEVFICTTPLLKYDHCVGEKYRWVEQNLGPEFVERIILTRDKTVVMGDLLIDDKDNI
              ▲                            ▲
hdNT-2  181  TGAEPTPSWEHVLFTACHNQHLQLQPPRRRLHSWADDWKAILDSKRPC
mdNT-2  173  TGAEPHPSWEHILFTSCHNYHLQLQPPRRRLHSWADDWKAILDSKRLR
hdNT-1  150  RGQEETPSWEHILFTCCHNRHLVLPPTRRRLLSWSDNWREILDSKRGAAQRE
mdNT-1  152  QGLEETPSWEHILFTCCHNQHLALPPTRRRLLSWSDNWRGIIESKRASL
              ▲
```
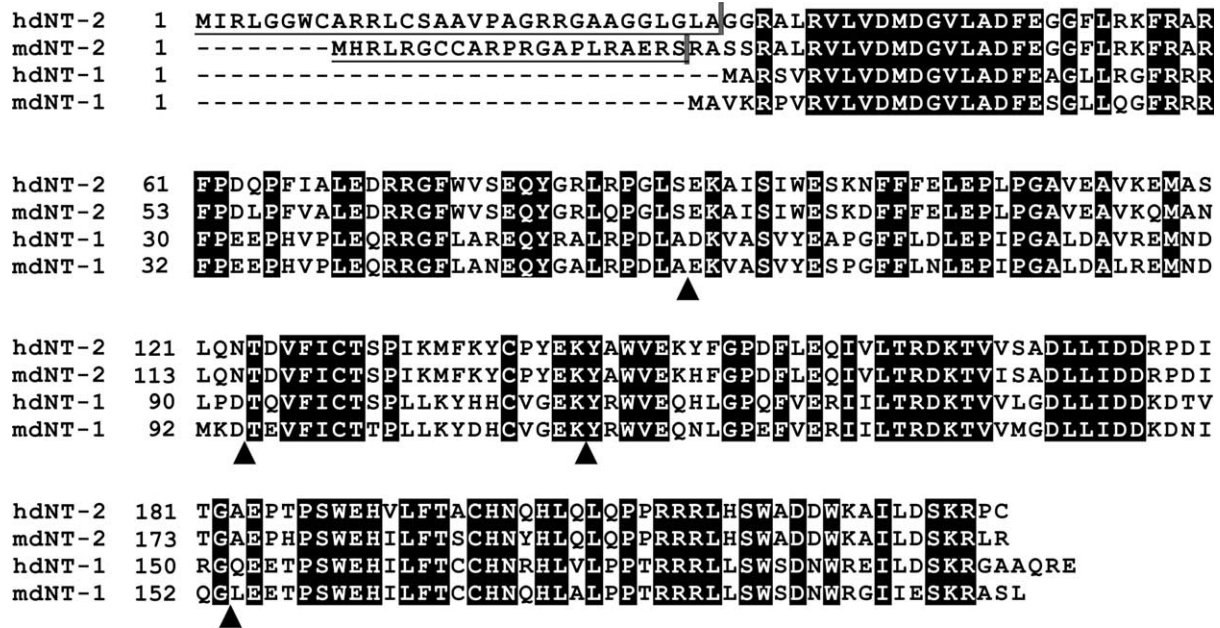
Fig. 1. Multiple alignment of the amino acid sequences of mouse (m) dNT-2 and dNT-1 with the human (h) counterparts. Identical amino acids in all four sequences are boxed. Filled triangles indicate the conserved exons borders in the four sequences. Underlined peptides are the putative leader sequences for mitochondrial import and bars indicate the proposed sites for cleavage of the pre-proteins during import.

also in the human dNT-2 and in two other recently described mitochondrial proteins (Spyrou et al., 1997; Lee et al., 1999). The enzymatic activity of the recombinant murine dNT-2 starting immediately downstream from the putative cleavage site (see below) strongly suggests that this is indeed the correct targeting signal.

3.2. Identification of the genomic sequences of the murine deoxyribonucleotidase genes

We identified genomic sequences homologous to the cDNAs of both mouse deoxyribonucleotidases. The GenBank clone AC068808 contained the *Nt5m* gene and clone AL645470 the *Nt5c* gene. The general structures of the two genes were similar and consisted, like the human counterparts (Rampazzo et al., 2000b), of five exons and four introns with borders located at identical exon sites (indicated by triangles in Fig. 1). The mouse *Nt5m* gene has much longer introns than the mouse *Nt5c* gene and covers 28.4 kb vs. the 1.4 kb of the latter (the corresponding figures are 44.3 and 1.5 kb for the human genes (Rampazzo et al., 2000b)). Human *NT5M* maps on chromosome 17p11.2, about 22 kb apart from the gene for subunit 3 of Cop 9 (*COPS3*) (Potocki et al., 1999) that runs in the opposite orientation. A comparable situation occurs in the mouse genome, where *Nt5m* is separated by 8.3 kb from nt 1 of the coding sequence of Cop 9 subunit 3 (*Cops3*). The two mitochondrial deoxyribonucleotidases are similar also at the cDNA level: they have longer 3'-UTRs than the dNT-1 cDNAs, and have sub-optimal Kozak sequences (Kozak, 1996) whereas the dNT-1 cDNAs conform to the Kozak rule and contain a G both at -3 and at $+1$ relative to the

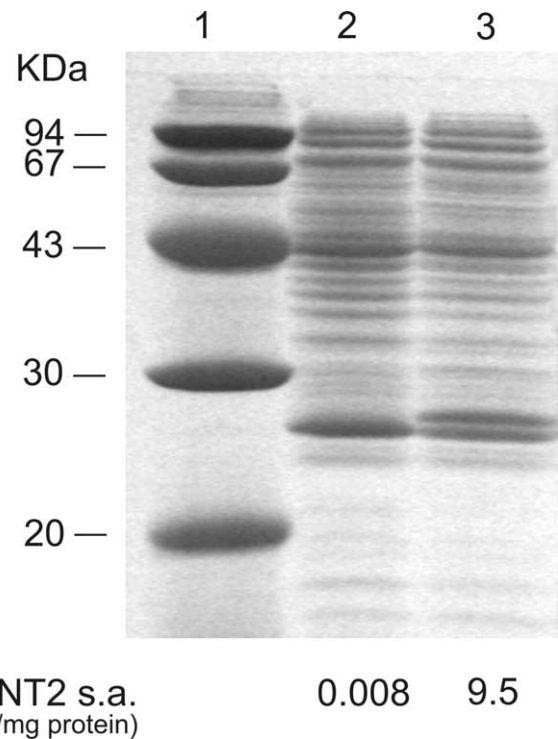initiation codon. These two features of the dNT-2 cDNAs together with some conserved intronic elements (see

Fig. 2. Overexpression of mouse dNT-2 in *E. coli*. SDS gel analysis of buffer extracts (5 μg protein) of bacteria transformed with plasmid p2M-mdNT-2 and induced with IPTG for 3 h (lane 3) or non induced (lane 2). Lane 1: size markers. Enzyme specific activity (s.a.) measured in the extracts is indicated at the bottom of the gel.

Section 3.5) may contribute to control the expression of mitochondrial deoxyribonucleotidases.

3.3. Expression of mouse dNT-2 in E. coli

We subcloned the portion of the mouse dNT-2 cDNA coding for the putative processed protein, i.e. the polypep-tide starting from R 22 with an additional 5′ ATG into an inducible vector and expressed it in *E. coli*. A prominent band of about 26 kDa was present in the electrophoretic pattern of protein extracts obtained from induced bacterial cultures but was absent from control cultures (Fig. 2). The specific activity of dNT-2 increased from 0.008 U/mg protein in the control to 9.5 in the induced bacteria. We
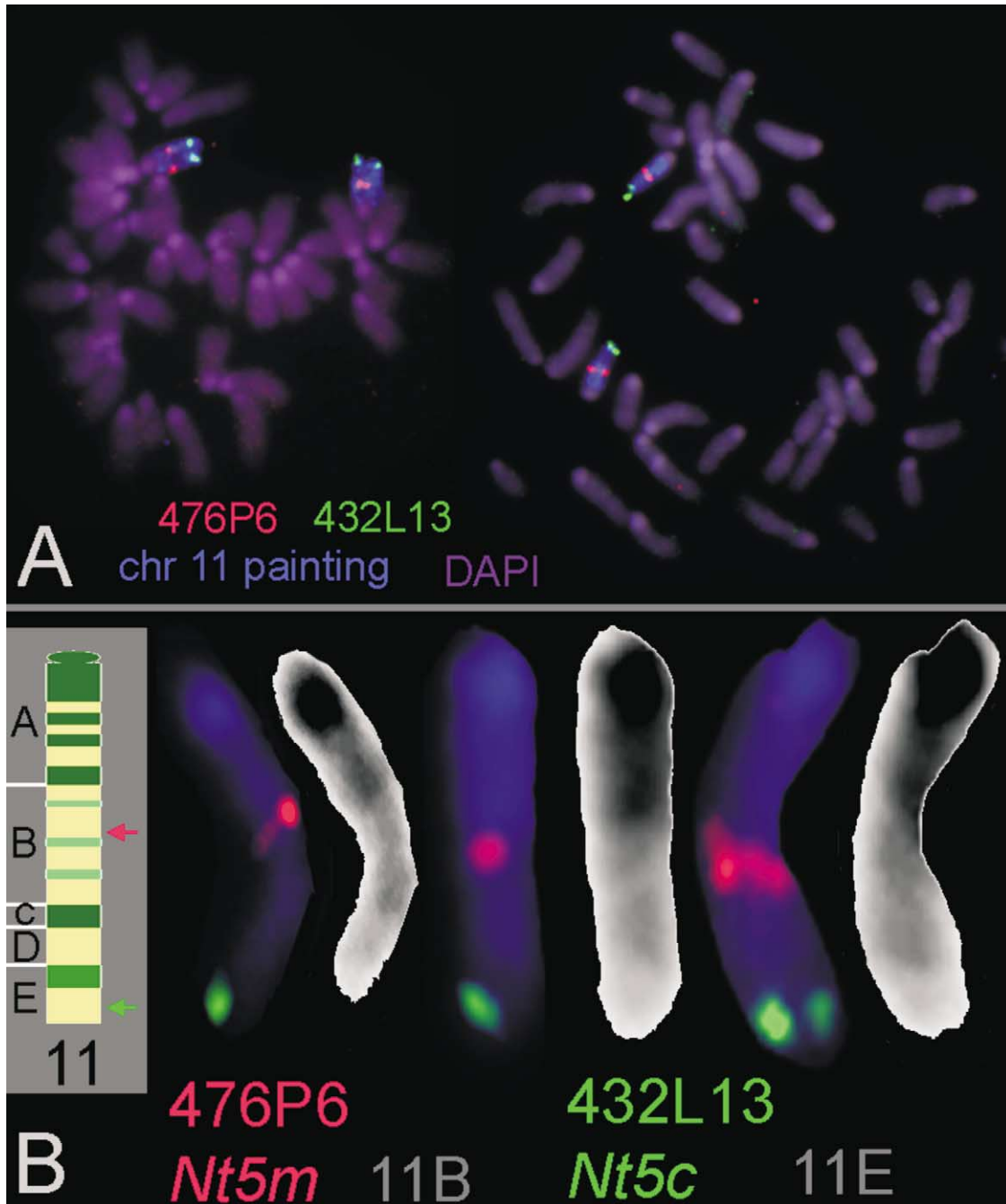

Fig. 3. Chromosomal localization of *Nt5c* (green: probe 432L13) and *Nt5m* (red: probe 476P6) by fluorescent in situ hybridization (FISH). Colors were artificially transformed as indicated. (A) Two metaphases with the FISH signals on mouse chromosomes 11 (blue: chromosome 11 specific painting probe). (B) Localization of probe 476P6 at 11B and probe 432L13 at 11E. Left: chromosome 11 ideogram with the position of two probes (arrows). The chromosomes are shown in two images. First: DAPI stained chromosome (blue) with the signals. Second: black and white inverted image of DAPI staining for probe assignment to chromosomal band (gray).

conclude that the cloned cDNA codes for a deoxyribonucleotidase active on dUMP that is the murine ortholog of the human dNT-2 and that during import into mitochondria the cleavage of the pre-protein most probably occurs at the site indicated in Fig. 1.

3.4. Characterization of genomic clones and chromosomal assignment of the murine Nt5c and Nt5m genes

In order to obtain genomic probes suitable to identify the localization of the two deoxyribonucleotidase genes on mouse chromosomes, we used cDNA fragments as probes to screen a mouse PAC library by filter hybridization. Six and four PAC clones were isolated for the murine *Nt5c* and *Nt5m* genes, respectively (432L13, 337F17, 346B20, 396H21, 446C13, 499J13 and 476P6, 552F24, 598E20, 668A16). DNA from each of these clones was isolated. Analysis by *Eco*RI restriction endonuclease digestion revealed a distinct pattern of restriction fragments for the two pools of clones, which confirmed that they are derived from two distinct loci in the mouse genome (details not shown). One PAC clone for each locus (432L13 and 476P6) was selected to determine the chromosomal localization of both genes using fluorescent in situ hybridization. The FISH signals for two PACs, representing *Nt5c* and

Nt5m genes were both detected on mouse chromosome 11. PAC 432L13 (*Nt5c* gene) was localized at band 11E and PAC 476P6 (*Nt5m* gene) at band 11B (Fig. 3). The human *NT5C* gene maps on chromosome 17q25 while *NT5M* maps at 17p11.2 in the critical region deleted in Smith–Magenis syndrome patients (Chen et al., 1996), which is located in the region showing synteny with the part of murine chromosome 11 containing the mouse *shaker-2* critical region (Probst et al., 1999). Our FISH results with the *Nt5m* specific probe demonstrate that the *Nt5m* gene is indeed located on chromosome 11.

3.5. Evolutionary analyses of loci for deoxyribonucleotidase 1 and 2 genes

Dot-plot and PIP (percent identity plot) analyses of the genes for the cytosolic and mitochondrial deoxyribonucleotidases revealed strong conservation between human and mouse genes. This conservation covered the entire 500 kb of studied sequences surrounding the genes (see Figs. 4 and 5). As expected, a strong conservation of genomic structure and exonic sequences was seen for both pairs of orthologous genes (*NT5C* and *Nt5c* as well as *NT5M* and *Nt5m*). Similarly, as shown in Figs. 4 and 5, the conservation can also be clearly seen for paralogs. For instance, the conservation between human *NT5C* and both human *NT5M* and mouse
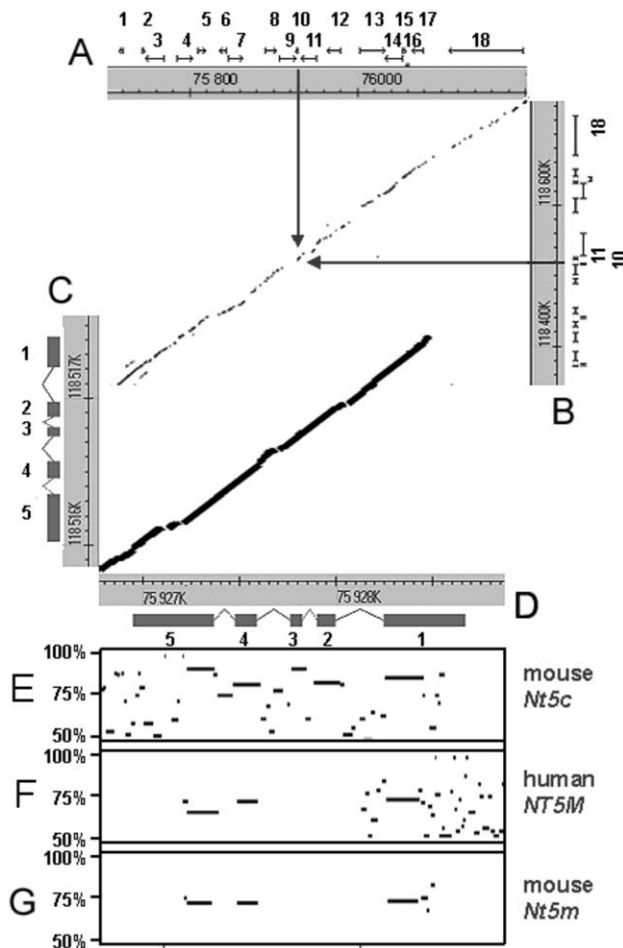

Fig. 4. Evolutionary conservation of the *NT5C* and *Nt5c* loci based on dot-plot and PIP (percent identity plot) analyses. Dot-plot analyses between the human 17q25 and mouse chromosome 11E sequences are shown on two levels. Large-scale analysis of 500 kb sequence between the two loci compares human sequence in horizontal axis (A) and mouse sequence in vertical axis (B). Genomic sequences are derived from the Celera's data base (scaffold GA_x9V1B8R:6500001..7413081 and scaffold GA_x5J8B7W64TB:1..1000000, respectively). The dot-plot output is shown in small dot format and positions of the *NT5C* and *Nt5c* genes are indicated by two long arrows. Numbers shown in panel (A) above arrows represent 18 genes from *Homo sapiens*, which are classified in the Celera data base as 'expert reviewed' or processed through 'Celera computational annotation (Otto)'. The following numbers correspond to the human genes characterized in the public data bases: 5, *ICT1* (acc. no. NM_001545); 8, *SLC16A5* (acc. no. NM_004695); 9, *FLJ22160* (acc. no. NM_024585); 10, *NT5C* (dNT-1) (acc. no. NM_014595); 11, *HN1* (acc. no. NM_016185); 12, *SMT3H2* (acc. no. NM_006937); 13, *FLJ12549* (acc. no. NM_024844); 15, *MRPS7* (acc. no. NM_015971); 16, *AD023* (acc. no. NM_020679); 17, *MUP1* (acc. no. NM_021734); 18, *GRB2* (acc. no. NM_002086). Scale in kb is corresponding to the position of the *NT5C* and *Nt5c* loci in the Celera data base on human chromosome 17 and mouse chromosome 11. In panel (B), number 10, 11 and 18 represent the only three genes which are also characterized in the mouse: *Nt5c* (NM_015807), *Hn1* (NM_008258) and *Grb2* (NM_008163), respectively. High-resolution dot-plot analysis (thick line) of approximately 1.5 kb, encompassing the *NT5C* and *Nt5c* genes only, between human sequence in horizontal axis (D) and mouse sequence in vertical axis (C). Panels (C) and (D) show also the genomic structure of the mouse *Nt5c* gene and the human *NT5C* gene, respectively, with five exons encompassing less than 1.5 kb. Panels (E–G) show results from three pair-wise PIP analyses of human genomic sequence for the *NT5C* gene (reference sequence) and the mouse *Nt5c* gene (E), the human *NT5M* gene (F) and the mouse *Nt5m* gene (G). Degree of identity scale between two nucleotide sequences (50–100%) is indicated on the left-hand side of each plot.
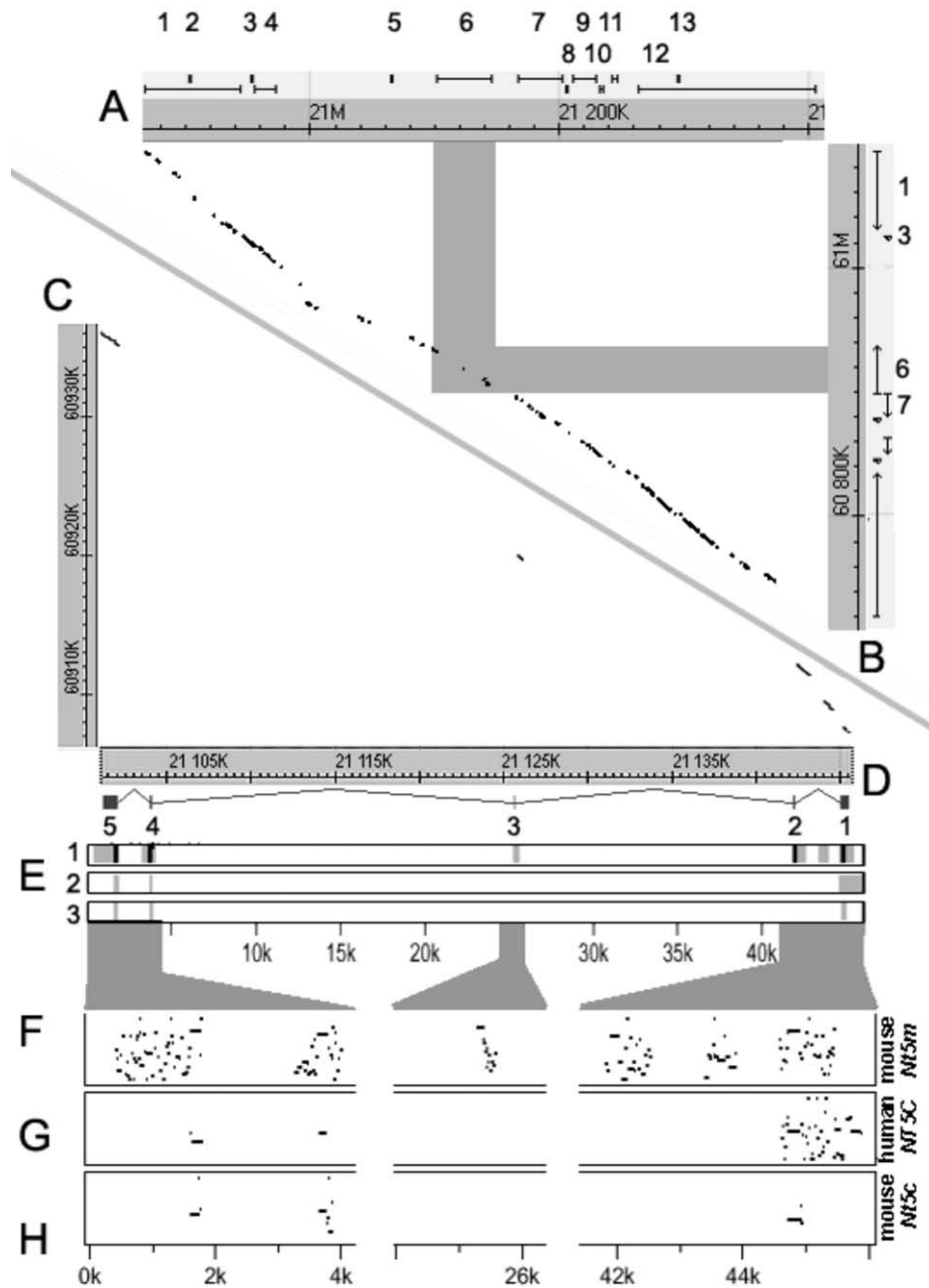
Fig. 5. Evolutionary conservation of the *NT5M* and *Nt5m* loci based on dot-plot and PIP (percent identity plot) analyses. Dot-plot analyses between the human 17p11.2 and mouse chromosome 11B sequences are shown on two levels. Large-scale analysis of 500 kb sequence between the two loci compares human sequence in horizontal axis (A) and mouse sequence in vertical axis (B). Genomic sequences are derived from the Celera data base (scaffold GA_x54K-RE9GCM9:500001..1500000 and scaffold GA_x5J8B7W3UM0:1..1000000, respectively). The positions of the *NT5M* and *Nt5m* genes are indicated by two gray bands. Numbers shown in panel (A) above horizontal lines represent 13 genes from *Homo sapiens*, which are classified in the Celera data base as 'expert reviewed' or 'processed through Celera computational annotation (Otto)'. The following numbers correspond to the human genes characterized in the public data bases: 1, *PEMT* (acc. no. BC000557); 2, *FLJ23022* (acc. no. NM_025051); 3, *LOC51655* (acc. no. NM_016084); 4, *FLJ10193* (acc. no. NM_018019); 5, *RPL13* (acc. no. BC000851); 6, *NT5M* (acc. no. NM_020201); 7, *COPS3* (acc. no. NM_003653). Scale in kb is corresponding to the position of the *NT5M* and *Nt5m* loci in the Celera data base of human chromosome 17 and mouse chromosome 11. In panel (B), numbers 1, 3 and 7 represent the only three genes which are also characterized in the mouse: *Pemt* (acc. no. NM_008819), *Rasd1* (acc. no. NM_009026) and *Cops3* (acc. no. NM_011991), respectively. Number 6 represents the gene processed through 'Celera computational annotation (Otto)' and corresponds to *Nt5m* (acc. no. BK 000190). High-resolution dot-plot analysis is displayed below the gray diagonal line. It shows comparison of approximately 46 kb, encompassing the *NT5M* and *Nt5m* genes only, between human sequence in horizontal axis (D) and mouse sequence in vertical axis (C). Panel (D) shows also genomic structure of the *NT5M* gene with five numbered exons. Panel (E) shows results from three pair-wise PIP analyses of human genomic sequence for the *NT5M* gene (reference sequence) and the mouse *Nt5m* gene (E1), the human *NT5C* gene (E2) and the mouse *Nt5c* gene (E3). The level of conservation between genomic sequences is represented by black bars (identity of nucleotide sequence of above 75%) and gray bars (identity of nucleotide sequence between 50 and 75%). Panels (F–H) represent a detailed view of six islands which showed conservation above 50% nucleotide sequence identity, and these correspond to pair-wise comparisons displayed in E1, E2 and E3, respectively.

Nt5m is clearly evident for three out five exons. Exons 2 and 3 are divergent in the paralogous genes. Conservation can also be seen for intronic sequences between both orthologs and paralogs. For instance, a clear island of conservation between genomic sequence within intron 1 of human *NT5M* and mouse *Nt5m* can be observed. Such conserved islands provide an opportunity for future characterization of elements which regulate gene transcription. Indeed, comparison of genomic sequences between species is emerging as a powerful tool for exploring unknown complexity of genome anatomy (Onyango et al., 2000).

3.6. Conclusions

(1) We have cloned the cDNA of the mouse mitochondrial deoxyribonucleotidase. The deduced amino acid sequence shows a high degree of similarity to the human dNT-2 and structural features for mitochondrial import. When expressed in bacteria in the processed form putatively present in mitochondria, the recombinant protein was a deoxyribonucleotidase actively dephosphorylating dUMP, the optimal substrate of human dNT-2.

(2) We have identified and annotated the genomic sequences of the murine deoxyribonucleotidases in the GenBank data base and isolated the corresponding genomic DNAs from a PAC library. Two genomic clones were used as probes for chromosomal mapping of the two mouse genes by FISH.

(3) *Nt5c* maps at 11E and *Nt5m* at 11B, in the *shaker-2* region, the mouse counterpart of the Smith–Magenis syndrome region on human chromosome 17. The Smith–Magenis syndrome is a multiple congenital anomalies syndrome associated with the deletion of a region covering different genes. It is still unclear if it arises from the loss of one critical gene or from a dosage effect involving several linked genes. Targeted disruption of *Nt5m* may reveal if dNT-2 deficiency contributes to the phenotype. We are currently using the information on the murine *Nt5m* obtained in the present study to prepare a knock-out model of such a situation.

(4) Evolutionary analyses of the mouse and human loci for the two deoxyribonucleotidases revealed strong conservation among the four genes. In particular, PIP analysis revealed that *NT5M* and *Nt5m* contain several islands of remarkable conservation within introns that may represent control elements. Future work will test this hypothesis.

Acknowledgements

References

Allegrini, S., Scaloni, A., Ferrara, L., Pesi, R., Pinna, P., Sgarrella, F., Camici, M., Eriksson, S., Tozzi, M.G., 2001. Bovine cytosolic $5'$-nucleotidase acts through the formation of an aspartate 52-phosphoenzyme intermediate. J. Biol. Chem. 276, 33526–33532.

Bianchi, V., Pontis, E., Reichard, P., 1986. Interrelations between substrate cycles and de novo synthesis of pyrimidine deoxyribonucleoside triphosphates in 3T6 cells. Proc. Natl. Acad. Sci. USA 83, 986–990.

Branda, S.S., Isaya, G., 1995. Prediction and identification of new natural substrates of the yeast mitochondrial intermediate peptidase. J. Biol. Chem. 270, 27366–27373.

Chen, K.S., Potocki, L., Lupski, J.R., 1996. The Smith–Magenis syndrome [del(17)p11.2]: Clinical review and molecular advances. Ment. Retard. Dev. Disabil. Res. Rev. 2, 122–129.

Claros, M.G., Vincens, P., 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 779–786.

Fedorova, L., Kost-Alimova, M., Gizatullin, R.Z., Alimov, A., Zabarovska, V.I., Szeles, A., Protopopov, A.I., Vorobieva, N.V., Kashuba, V.I., Klein, G., Zelenin, A.V., Sheer, D., Zabarovsky, E.R., 1997. Assignment and ordering of twenty-three unique NotI-linking clones containing expressed genes including the guanosine $5'$-monophosphate synthetase gene to human chromosome 3. Eur. J. Hum. Genet. 5, 110–116.

Feinberg, A.P., Vogelstein, B., 1984. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity (Addendum). Anal. Biochem. 137, 266–267.

Gazziola, C., Ferraro, P., Moras, M., Reichard, P., Bianchi, V., 2001. Cytosolic high K(m) $5'$-nucleotidase and $5'(3')$-deoxyribonucleotidase in substrate cycles involved in nucleotide metabolism. J. Biol. Chem. 276, 6185–6190.

Itoh, R., 1993. IMP-GMP $5'$-nucleotidase. Comp. Biochem. Physiol. B 105, 13–19.

Johansson, M., Karlsson, A., 1997. Cloning of the cDNA and chromosome localization of the gene for human thymidine kinase 2. J. Biol. Chem. 272, 8454–8458.

Kozak, M., 1996. Interpreting cDNA sequences: some insights from studies on translation. Mamm. Genome. 7, 563–574.

Lee, S.R., Kim, J.R., Kwon, K.S., Yoon, H.W., Levine, R.L., Ginsburg, A., Rhee, S.G., 1999. Molecular cloning and characterization of a mitochondrial selenocysteine-containing thioredoxin reductase from rat liver. J. Biol. Chem. 274, 4722–4734.

Neupert, W., 1997. Protein import into mitochondria. Annu. Rev. Biochem. 66, 863–917.

Onyango, P., Miller, W., Lehoczky, J., Leung, C.T., Birren, B., Wheelan, S., Dewar, K., Feinberg, A.P., 2000. Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. Genome Res. 10, 1697–1710.

Osoegawa, K., Tateno, M., Woon, P.Y., Frengen, E., Mammoser, A.G., Catanese, J.J., Hayashizaki, Y., de Jong, P.J., 2000. Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. Genome Res. 10, 116–128.

Paglia, D.E., Valentine, W.N., Brockway, R.A., 1984. Identification of thymidine nucleotidase and deoxyribonucleotidase activities among normal isozymes of $5'$-nucleotidase in human erythrocytes. Proc. Natl. Acad. Sci. USA 81, 588–592.

Potocki, L., Chen, K.S., Lupski, J.R., 1999. Subunit 3 of the COP9 signal transduction complex is conserved from plants to humans and maps within the Smith–Magenis syndrome critical region in 17p11.2. Genomics 57, 180–182.

Probst, F.J., Chen, K.S., Zhao, Q., Wang, A., Friedman, T.B., Lupski, J.R., Camper, S.A., 1999. A physical map of the mouse shaker-2 region contains many of the genes commonly deleted in Smith–Magenis syndrome (del17p11.2p11.2). Genomics 55, 348–352.

Rampazzo, C., Johansson, M., Gallinaro, L., Ferraro, P., Hellman, U., Karlsson, A., Reichard, P., Bianchi, V., 2000a. Mammalian $5'(3')$-deoxyribonucleotidase, cDNA cloning, and overexpression of the enzyme in *Escherichia coli* and mammalian cells. J. Biol. Chem. 275, 5409–5415.

Rampazzo, C., Gallinaro, L., Milanesi, E., Frigimelica, E., Reichard, P., Bianchi, V., 2000b. A deoxyribonucleotidase in mitochondria: involvement in regulation of dNTP pools and possible link to genetic disease. Proc. Natl. Acad. Sci. USA 97, 8239–8244.

Reichard, P., 1988. Interactions between deoxyribonucleotide and DNA synthesis. Annu. Rev. Biochem. 57, 349–374.

Resta, R., Yamashita, Y., Thompson, L.F., 1998. Ecto-enzyme and signaling functions of lymphocyte CD73. Immunol. Rev. 161, 95–109.

Sala-Newby, G.B., Skladanowski, A.C., Newby, A.C., Makarewicz, W., 1999. The mechanism of adenosine formation in cells. Cloning of cytosolic $5'$-nucleotidase-I. J. Biol. Chem. 274, 17789–17793.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schwartz, S., Zhang, Z., Frazer, K.A., Smith, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., Miller, W., 2000. PipMaker – a web server for aligning two genomic DNA sequences. Genome Res. 10, 577–586.

Shadel, G.S., Clayton, D.A., 1997. Mitochondrial DNA maintenance in vertebrates. Annu. Rev. Biochem. 66, 409–435.

Spyrou, G., Enmark, E., Miranda-Vizuete, A., Gustafsson, J.Å., 1997. Cloning and expression of a novel mammalian thioredoxin. J. Biol. Chem. 272, 2936–2941.