Identification and characterization of the t(w73) candidate gene Ortc3
Verhaagh, S.F.M.J.

Citation for published version (APA):
Verhaagh, S. F. M. J. (2001). Identification and characterization of the t(w73) candidate gene Ortc3

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 2

Cloning of the mouse and human solute carrier 22a3 (Slc22a3/SLC22A3) identifies a conserved cluster of three organic cation transporters on mouse chromosome 17 and human 6q26–q27

S. Verhaagh, N. Schweifer, D.P. Barlow, and R. Zwart

Cloning of the mouse and human solute carrier 22a3 (Slc22a3/SLC22A3) identifies a conserved cluster of three organic cation transporters on mouse chromosome 17 and human 6q26–q27

S. Verhaagh, N. Schweifer¹, D.P. Barlow, and R. Zwart

Department of Molecular Genetics, The Netherlands Cancer Institute, Amsterdam, The Netherlands
¹ Present address: Department of Molecular Biology, Boehringer Ingelheim R&D Vienna, Vienna, Austria

Here we report the isolation of the mouse and human solute carrier genes Slc22a3/SLC22A3. Slc22a3 is specifically expressed in placenta, but the levels of expression decline towards the end of gestation. A BAC contig spanning the mouse Slc22a3 gene was constructed, and Slc22a3 was mapped between the Igf2r and Plg genes in close association with two additional members of the Slc22a gene family, Slc22a1 and Slc22a2. A partial cDNA sequence of the SLC22A3 gene was reconstituted from sequenced EST clones. SLC22A3 is expressed in first-trimester and term placenta, but also in skeletal muscle, prostate, aorta, liver, fetal lung, salivary gland, and adrenal gland. Using a somatic cell hybrid panel and a human YAC clone, SLC22A3 was mapped to the syntenic region on human chromosome 6q26–q27, between the IGF2R and APO(a)-like genes. SLC22A1 and SLC22A2 localized to the same locus, demonstrating the conservation of the close physical linkage of these three organic cation transporter genes in mouse and humans.

Keywords: t complex, t⁷ underscores, organic cation transporter (Orct), placenta, mouse chromosome 17, human 6q26–q27
Chapter 2

**Introduction**

$\tau$ haplotypes are structurally variant forms of the proximal part of mouse chromosome 17, carrying four large genomic inversions over a total length of 20-30 cM. This results in a severe suppression of recombination that has been instrumental in the identification of multiple genetic loci, which are involved in a variety of developmental defects (1). Two mutations have been mapped to a 500-kb region that is deleted in the partial $\tau$ haplotype allele $\tau^{wLub2}$ (2, 3). One is the T-associated maternal effect ($Tme$), which results in a late embryonic lethality of heterozygotes when the $\tau^{wLub2}$ allele is inherited from the mother, but not from the father (4). This effect has been assigned to the absence of the imprinted $Ig/2r$ (insulin-like growth factor receptor, also known as the cation independent mannose-6-phosphate receptor) gene, which was mapped within the $\tau^{wLub2}$ deletion (5). A second mutation that cannot be complemented by the $\tau^{wLub2}$ allele is located on the mutant $\tau$ haplotype allele $\tau^{w73}$. Mice homozygous for the $\tau^{w73}$ mutation die very early in embryonic development. At 6 days post coitum (dpc), these embryos show a retardation in growth and fail to form an ectoplacental cone (6).

Previously, the C3 YAC contig was constructed spanning the entire $\tau^{wLub2}$ deletion region (7). This contig was used to map the relative positions of genes known to reside within the $\tau^{wLub2}$ deletion with respect to the $Ig/2r$ gene and investigate their imprint status. Furthermore, it served as a probe in direct cDNA library hybridization experiments to identify new genes located within the $\tau^{wLub2}$ deletion. This resulted in the isolation of a gene coding for a polyspecific transmembrane transporter protein (8). This gene was initially called $Lxl$ because of its liver-specific expression and was the mouse homologue of the rat organic cation transporter 1 ($Oct1$) gene (9). Rat $Oct1$ was identified by functional expression cloning as a mediator of organic cation uptake over the basolateral membrane of the renal proximal tubules, as well as hepatocytes. Subsequently, an increasing number of organic cation transporters have been identified. The mouse homologue of the rat $Oct2$ gene has recently been deposited in the database, and the gene was assigned to chromosome 17, like mouse $Oct1$ (GenBank Accession No. AJ006036; Mooslehner, 1998). Organic cation transporters are predicted to contain a 12-transmembrane domain structure, and their kidney- and liver-specific expression patterns imply an essential role in the organ-specific transport of various molecules (10). Based on the homology in protein structure, the organic cation transporters have recently been assigned to the family of solute carrier proteins, and the $Oct1$ and $Oct2$ genes have subsequently been renamed $Slc22a1$ and $Slc22a2$, for solute carrier family 22a, members 1 and 2 (http://www.informatics.jax.org).

We report here the identification of the mouse and human homologues of a new member of the solute carrier family, $Slc22a3/SLC22A3$, which is homologous to the previously identified $Slc22a1$ and $Slc22a2$ genes. Mouse $Slc22a3$ was isolated from the $\tau^{wLub2}$ deletion region by Cpg island mapping and exon trapping techniques. In contrast to mouse $Slc22a1$ and 2, no expression is detected in liver or spleen, but high expression is observed in placenta that decreases dramatically towards the end of gestation. Analysis of a BAC (bacterial artificial chromosome) contig showed that all three $Slc22a$ genes were clustered within a 300-kb genomic region between the $Ig/2r$ and $Plg$ (plasminogen) genes on mouse chromosome 17. The human $SLC22A3$ gene was isolated by EST (expressed sequence tag) database screening, and was found to be more widely expressed than its mouse homologue. Analysis of a somatic cell hybrid panel mapped $SLC22A3$ to human chromosome 6, which contains a region of synteny with the mouse $\tau^{wLub2}$ region at 6q26-q27. Analysis of a human YAC clone showed that
Conserved Slc22a gene cluster in mouse and human

SLC22A3 mapped between the IGF2R and the APO(a)-like genes in a conserved cluster with SLC22A1 and SLC22A2.

Results

Identification of the mouse solute carrier 22a3 (Slc22a3) gene

In an attempt to identify new genes within the ^p^LUB2 deletion region, CpG island mapping was performed. CpG islands are short stretches of nonmethylated, CpG-rich sequences within the mammalian genome. They are associated with the 5' end of approximately 50% of all expressed genes in the mouse (11) and are therefore a powerful tool in the identification of previously unknown genes within a genomic locus. Within the P1 phage 55D4, which spans the proximal breakpoint of the ^p^LUB2 deletion (Fig. 1; (7)), a cluster of BssHII and SacII restriction sites was identified. DNA sequence analysis of 1790 bp of a 4.3-kb BamHI genomic subclone spanning the BssHII and SacII restriction sites (GenBank Accession No. AF078748) revealed a 605-bp region that contained a CpG dinucleotide content equivalent to the GpC dinucleotide number (Fig. 1). Restriction enzyme analyses with HpaII andMspI showed that this region was free of methylation (data not shown).

To determine whether the CpG island is indeed associated with the 5' end of an expressed gene, a 740-bp PstI fragment spanning the CpG island was used as a probe on a Northern blot containing adult tissues and 12.5 dpc mouse embryo and placenta samples (Fig. 3A). A 3.5-kb mRNA was detected in 12.5 dpc placental tissue but not in the embryo. Upon longer exposure, low levels of expression could be detected in adult kidney and brain. In addition, RT-PCR followed by hybridization with an internal oligonucleotide did not detect any expression in liver or spleen (data not shown).

In addition to the identification of a CpG island, exon trapping of BamHI fragments of P1 phage 55D4 (see Materials and Methods) resulted in the isolation of one putative exon

![Figure 1](image_url)

**Figure 1** Schematic representation of the 500-kb ^p^LUB2 deletion region (thick line) showing the relative positions of the Sod2, Mas, Igf2r, and Slc22a1 genes. Plg is located approximately 40 kb upstream of the ^p^LUB2 proximal breakpoint, whereas Sod2 lies 20 kb upstream of the distal breakpoint. The position of P1 phage 55D4, which spans the proximal end of the ^p^LUB2 deletion, is indicated. The gray oval represents a potential CpG island that was identified in P1 55D4 using BssHII and SacII restriction enzymes. In the map plot below, the presence of CpG and GpC dinucleotide pairs is shown, as is the location of the HpaII/MspI, SacII, MseI, and BssHII restriction sites, which were identified upon DNA sequence analysis of the first 1790 bp of a 4.3-kb BamHI subclone from P1 55D4 containing the CpG island. A 605-bp region was identified showing a high CpG dinucleotide density, which reached almost a 1:1 ratio with GpC dinucleotide pairs.

37
Conserved Slc22a gene cluster in mouse and human

**Figure 2** The gene associated with the CpG island is a member of the solute carrier 22a family of organic cation transporter proteins.

A cDNA sequence and amino acid translation of the mouse Slc22a3 gene. Here, only the 5' untranslated and coding regions are shown. The complete cDNA sequence is available under GenBank Accession No. AF078750. Nucleotide numbering is indicated at the left and amino acid numbering at the right. The total length of the mouse Slc22a3 cDNA is 3499 bp, containing a 1463 bp 3' untranslated region. The entire CpG island is contained within the 5' end of the Slc22a3 cDNA (nucleotide position 183-787) and includes the coding sequence for the first 136 amino acids. The amino acid translation is indicated in single capitals. The stop codon is indicated by an asterisk. The protein encodes 12 putative transmembrane domains, which are underlined. Possible N-glycosylation sites (printed in bold) are present at amino acid positions 72, 99, 114, 199, and 317. Potential PKA and PKC phosphorylation sites are double-underlined and located at position numbers 346 and 544 and 286, 292, and 459, respectively. B Multiple sequence alignment of the organic cation transporters. The PileUp multiple sequence alignment program was run with the amino acid sequences of Slc22a3 and 15 previously reported organic cation transporters: Orct1 (Nkt; (12)); Slc22a12 (Rst; (13)); rat Nlt (14); rat Slc22a3 (Orct3; (15)); mouse Slc22a2 (GenBank Accession No. AJ006036; Mooslehner, 1998); rat Slc22a2 (16); pig Slc22a2 (17); human SLC22A1 and 2 (18); mouse Slc22a1 (8); rat Slc22a1 (9); dmOrct1 and 2 (19); human SLC22A4 (OCTN1; (20)); and human SLC22A5 (OCTN2; (21)). The dendogram reflects only the relative homologies in primary amino acid sequence, not a phylogenetic relationship. The gray box highlights the position of the Slc22a3 protein.

that recognized the same 3.5-kb placenta-specific mRNA in a Northern blot analysis (data not shown). Surprisingly, DNA sequence analysis of this trapped exon showed that it was homologous to, but not identical with, the Slc22a1 gene which was previously isolated from the p<sup>\\text{Lub2}</sup> deletion. A complete cDNA containing the CpG island and trapped exon sequences was isolated from a 13.5 dpc placental cDNA phage library (gift from Drs S. Varmuza and J. Rossant). The total cDNA length is 3499 bp, containing an open reading frame of 1653 nucleotides, which codes for a protein of 551 amino acids (Fig. 2A; GenBank Accession No. AF078750). The CpG island sequence runs from nucleotides 183 to 787, whereas the trapped
Chapter 2

The exon is located between nucleotide numbers 888 through 1053. The BLASTX database search program (22) confirmed the identity of the newly isolated gene as a member of the family of organic cation transporters. Forty-six percent identity was found in the amino acid sequence with the mouse Slc22a1 and Slc22a2 genes. However, the protein is 98% homologous to the recently isolated rat Orct3 protein, now renamed Slc22a3, which was shown to transport a variety of organic cation molecules (15). We therefore propose that the gene associated with the CpG island close to the proximal breakpoint of \( t^{wab2} \) is the mouse homolog of the rat Slc22a3 gene.

The transmembrane prediction program Tmpred at ISREC (23) scores a strong preference for a putative 12-transmembrane domain (TMD) structure (underlined in Fig. 2A), with the NH\(_2\)- and C-terminal ends of the protein facing the cytoplasm. An alternative prediction would discard TMD11 and cause a switch of TMD1 to TMD10 of their inside-to-outside or outside-to-inside orientations. Running a Prosite Pattern Search (http://www2.ebi.ac.uk/ppsearch) reveals three possible N-glycosylation sites within the first extracellular loop of the protein (Fig. 2A), which are in accordance with the predicted N-glycosylation sites in other Slc22a proteins. However, mouse and rat Slc22a3 can be distinguished from the other Slc22a proteins by two additional N-glycosylation sites, which are positioned between TMD3 and TMD4 and between TMD6 and TMD7 (Asn-199 and Asn-317, respectively; Fig. 2A). Interestingly, in addition to the more generally conserved sites, one protein kinase A site (Thr-544) and one protein kinase C site (Thr-459) were found to be unique for the Slc22a3 protein (Fig. 2A). Although the predicted size of the Slc22a3 protein is not different from any of the other Slc22a proteins cloned thus far, its mRNA does contain an exceptionally long 3' untranslated region of 1463 nucleotides.

A Pileup multiple sequence alignment of the organic cation transporter proteins is depicted as a dendogram in Fig. 2B. It groups the Slc22a1, 2, and 3 proteins that have been isolated from human, rat, and mouse as the members with closest homology, although Slc22a1 and Slc22a2 are more closely related to each other than to Slc22a3. All other putative organic cation transporters appear more distantly related, and it is of interest to note that some of the proteins included in this alignment have been defined as an organic cation transporter based only on amino acid homology (Orct1, rat Nlt, and Slc22a2l2, as well as the dmOrct1 and 2 proteins identified in \textit{Drosophila melanogaster}), and some, in contrast to the prototype organic cation transporters, appear to function in an ATP-dependent manner (SLC22A4).

To delineate further the Slc22a3 expression pattern in placenta, both embryo mRNA and placenta mRNA were isolated at different stages of gestation. Figure 3B shows that a 3.5-kb mRNA is detected in placenta at all stages that were examined, but the levels of expression decline from 15.5 dpc onwards. Expression in the embryo could not be detected at 12.5 or 15.5 days of embryonic development (Fig. 3B). Thus, the Slc22a3 gene is a distinct, new member of the Slc22a gene family both in sequence and in expression distribution.

Three mouse Slc22a genes are closely linked within the proximal end of the \( t^{wab2} \) deletion on chromosome 17

The unexpected finding of a second Slc22a gene within the \( t^{wab2} \) deletion led us to investigate the genomic organization of these genes. Therefore, a BAC contig was isolated spanning a genomic region of approximately 300 kb that included the entire Slc22a1 gene and the 3' end of the \( Plg \) gene (Fig. 4). The Slc22a1 gene marks the distal end of the contig. It was previously shown that Slc22a1 gene maps in the \( t^{wab2} \) deletion, 20 kb downstream of the
Conserved Slc22a gene cluster in mouse and human

Figure 3 The CpG island is associated with a placental-specific gene. A Northern blot analysis of adult tissues and embryo and placenta. A 3.5-kb mRNA is detected only in placenta of 12.5 dpc, but not in embryos nor in any of the adult tissues tested. Actin was hybridized to the same blot as a quantitative loading control. B Northern blot analysis of the temporal pattern of expression in the maturing placenta. The 3.5-kb mRNA is detected in placenta from 12.5 dpc to 18.5 dpc. Expression declines from 15.5 dpc until birth, which is at 18-19 dpc in the FVB strain of mice used.

imprinted Igf2r gene (8) and Fig. 1). The proximal end of the BAC contig is marked by the 3' end of the Plg gene, which had been mapped 30-40 kb outside the inhib deletion (24). Using the 4.3-kb BamHI CpG island subclone as the 5' probe and a 806-bp EcoRI 3' cDNA fragment (see Materials and Methods) as the 3' probe, it was shown that the Slc22a3 gene spans a total genomic distance of 110 kb. It has the same transcriptional orientation as Slc22a1, and the two genes are separated by 140 kb. The 3' ends of the Slc22a3 and Plg genes were found to be in very close proximity to each other, as they were mapped to a single 2.2 kb HindIII genomic fragment (data not shown). Sequence analysis of this fragment revealed that the two genes did not overlap, but were separated by approximately 1 kb (data not shown). To investigate the possibility that the recently identified Slc22a2 gene, which was assigned to mouse chromosome 17 (GenBank Accession No. AJ006036, Mooslehner 1998), was located close to the Slc22a1 and Slc22a3 genes, we used Slc22a2 5' and 3' cDNA primers (see Materials and Methods). Slc22a2 does map to the BAC contig and is located in between the Slc22a1 and Slc22a3 genes in an inverse orientation (Fig. 4).

Isolation of the human SLC22A3 gene and conservation of its close linkage to SLC22A1 and
Chapter 2

2 on chromosome 6q26-q27

A search of the EST database with the complete mouse Slc22a3 cDNA sequence using BLASTN revealed two human EST clones from a fetal liver/spleen cDNA library (ID123234 and 127120; see Table 1). The EST clones were used to search the database for longer overlapping EST clones in an attempt to identify the complete SLC22A3 cDNA sequence. A total of 11 additional human SLC22A3 EST clones were isolated from diverse cDNA libraries (Table 1), spanning a total cDNA length of 1664 bp (GenBank Accession No. AF078749). The EST contig showed a homology of 68% with the mouse Slc22a3 cDNA sequence, which was greater than the homology to SLC22A1 and 2. This suggests that the identified ESTs constitute the 3' part of the human SLC22A3 gene. Furthermore, when EST clone ID 123234 was used as a probe on first-trimester placental and B-lymphoid cell line BJA-B RNA samples, a mRNA of approximately 3.6 kb, a size similar to that of Slc22a3, was identified in placenta (Fig. 5A). The diversity of cDNA library sources containing SLC22A3 EST clones suggests a more widespread expression in human than in mouse. To examine this more closely, the expression pattern of SLC22A3 in a wide range of tissues was analyzed using a normalized human RNA MasterBlot (Clontech). Strongest expression was observed in aorta, prostate, adrenal gland, salivary gland, skeletal muscle, liver, fetal lung, and term placenta, whereas no expression was detected in brain and spleen (Fig. 5B).

To determine the chromosomal localization of SLC22A3, a PCR analysis with primers specific for human SLC22A3 was performed on a human monochromosomal somatic cell hybrid DNA panel (HGMP; (26)). This experiment mapped SLC22A3 to human chromosome 6 (data not shown). The 6q26-q27 telomeric region of chromosome 6 is syntenic with the proximal part of mouse chromosome 17 containing the \( t^{Lub2} \) deletion region. Previously, a human 2-Mb human YAC contig had been isolated from this region, showing that the PLG/APO(a) genes lie close to IGF2R (27). The most distal YAC clone (204B2) of this contig was used to investigate whether the localization of Slc22a3 between the Igf2r and the Plg

![Figure 4](image)

**Figure 4** Three members of the Slc22a family, Slc22a1, Slc22a2, and Slc22a3, are physically linked on mouse chromosome 17.

A physical map of the Slc22a locus is shown at the top. BssHII (B) and Clal (C) sites are indicated. The position and orientation of the Slc22a1, 2, and 3 genes are indicated by arrows. For Plg, only the 3' end of the gene is indicated, which is separated by 1 kb from the Slc22a3 3' end. For Slc22a2, the exact size of the gene was not determined, and therefore the maximal genomic distance is indicated by a dashed arrow. The minimal distance between the Slc22a1 and Slc22a2 3' ends is 5 kb, whereas the Slc22a2 and Slc22a3 genes are separated by approximately 60 kb. The mouse Slc22a3 gene spans 110 kb. The BAC contig is shown below, spanning approximately 300 kb.
Conserved Slc22a gene cluster in mouse and human

Table 1 A wide range of tissues contain SLC22A3 ESTs

<table>
<thead>
<tr>
<th>IMAGE Clone ID Nos.</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>123234</td>
<td>Soares fetal liver/spleen</td>
</tr>
<tr>
<td>127120</td>
<td>Soares fetal liver/spleen</td>
</tr>
<tr>
<td>201787*</td>
<td>Soares fetal liver/spleen</td>
</tr>
<tr>
<td>436500*</td>
<td>Soares fetal liver/spleen</td>
</tr>
<tr>
<td>296332</td>
<td>Soares fetal liver/spleen</td>
</tr>
<tr>
<td>Fl-127D*</td>
<td>Fetal liver</td>
</tr>
<tr>
<td>1658458*</td>
<td>Soares total fetus</td>
</tr>
<tr>
<td>486785*</td>
<td>Soares pregnant uterus</td>
</tr>
<tr>
<td>795603*</td>
<td>Soares testis</td>
</tr>
<tr>
<td>1524902*</td>
<td>Soares testis/B-cell/fetal lung</td>
</tr>
<tr>
<td>913989*</td>
<td>NCI_CGAP normal prostate</td>
</tr>
<tr>
<td>1115897*</td>
<td>NCI_CGAP malignant prostate</td>
</tr>
<tr>
<td>1115553*</td>
<td>NCI_CGAP malignant prostate</td>
</tr>
</tbody>
</table>

Thirteen ESTs for the human SLC22A3 gene were identified in a database search, which are listed by their IMAGE Clone ID numbers. The ESTs were sequenced from a variety of cDNA libraries, as indicated in the right column. EST clones were obtained from the IMAGE Consortium (25).

* These ESTs show homology with human PLG exon 19 sequence.

genes is conserved in human. The top panel in Figure 6 shows a schematic diagram of the position of the previously mapped genes on YAC 204B2, which had been sized as 530 kb and mapped for BssHII and SalI restriction sites (27). The identity of YAC clone 204B2 was confirmed by hybridization with human IGF2R and APO(a) probes (data not shown). When YAC 204B2 was digested with SalI, SfiI, and BssHII, SLC22A3 recognized a 280-, a 470-, and a 180-kb fragment, respectively (Fig. 6). This result shows that SLC22A3 maps to YAC 204B2. Furthermore, the same fragments are recognized by APO(a)-like ((27) and data not shown), indicated that SLC22A3 is located on 6q26-q27 in close proximity to the APO(a)-like gene. The clustering of the mouse Slc22a3 and Slc22a1 and 2 genes in a chromosomal region that shows homology of synteny with human 6q26-q27 suggests that this cluster of Slc22a genes is conserved in human. In fact, it was recently reported that the human SLC22A1 and 2 genes map to 6q26 by DNA-FISH analysis (28), providing further support for this hypothesis. To determine the physical map for the human SLC22A genes, YAC 204B2 was hybridized with SLC22A1- and SLC22A2-specific probes (Fig. 6). In an SfiI digest, both SLC22A1 and 2 recognize the same 470-kb fragment that contains SLC22A3 and APO(a)-like, showing that in humans the clustering of the SLC22A genes is conserved and confined to a 470-kb region. However, SLC22A1 and 2 map to a different BssHII fragment of approximately 260 kb, showing that SLC22A1 and 2 are distal to SLC22A3 (Fig. 6). Finally, analysis of the SalI fragment pattern within the 260-kb BssHII fragment suggests that SLC22A1 is distal to SLC22A2. Given these data, we suggest the following gene order: telomere-IGF2R-SLC22A1-SLC22A2-SLC22A3-APO(a)-like-APO(a)-centromere.
Chapter 2

Figure 5 SLC22A3 gene expression. A Northern blot analysis of human SLC22A3 gene expression in placenta. A 3.6-kb mRNA is detected in first-trimester placenta (lane 1), but not in the human B-lymphoid cell line BJA-B (lane 2). The 18S and 28S ribosomal RNAs are indicated. B Normalized human RNA MasterBlot (Clontech), showing high expression in aorta (C2), skeletal muscle (C3), prostate (C7), adrenal gland (D5), salivary gland (D7), liver (E2), term placenta (F4), and fetal lung (G7). Moderate to low expression was detected in uterus (C6), ovary (D2), kidney (E1), lymph node (E7), lung (F2), trachea (F3), and fetal liver (G4), and no expression was seen in brain (A1-A7, B1-B6), pancreas (D3), pituitary gland (D4), spleen (E4, G5), peripheral leukocytes (E6), fetal brain (G1), fetal spleen (G5), and fetal thymus (G6).

Discussion

We have isolated the mouse solute carrier gene Slc22a3 in a search for genes located in the proximal part of the pterob deletion region. By Northern blot analysis, Slc22a3 shows a placenta-specific expression during embryonic development and low levels of expression in adult kidney and brain, but not liver. Expression in placenta was initially high, but decreased from 15.5 dpc until the end of gestation. In contrast, human SLC22A3 was expressed in first-trimester and term placenta, but was also present in a range of adult tissues including aorta, prostate, salivary gland, adrenal gland, fetal lung, skeletal muscle, and liver. Analysis of a BAC contig around the Slc22a3 genomic locus revealed a close physical linkage with two additional Slc22a family members, Slc22a1 and 2. All three Slc22a genes map between Igf2r and Plg on mouse chromosome 17. A similar analysis of a human YAC clone demonstrated that SLC22A3 was located together with SLC22A1 and 2 on chromosome 6q26-q27 between the IGF2R and PLG/APO(a) genes. This result shows that the organization of the Slc22a genes is conserved in mouse and humans on a chromosomal region with extensive synteny.

Placenta-specific expression of Slc22a3

The placenta is an essential organ of the post-implantation embryo for the exchange of nutrients and waste products between the growing fetus and the maternal environment. Glucose is the principal source of carbohydrates for the fetus, which is transported by the SLC2A1 (previously named GLUT) facilitative transporter protein (29). SLC2A1 shows structural homologies with the Slc22a proteins (8), and both the tissue-specific and developmental-
Conserved Slc22a gene cluster in mouse and human

![Restriction enzyme map of human SLC22A gene cluster](image)

**Figure 6** Organization of the human SLC22A gene cluster at chromosome 6q26-q27.

**Top** Representation of the restriction enzyme map as was determined from the 2-MB human YAC contig (27). The positions of the previously mapped IGF2R, APO(a)-like and APO(a) genes are indicated by the gray boxes. The boxes representing SLC22A1, 2, and 3 show the maximal genomic distance of the 3' ends of these genes. **Bottom:** Identical panels of SalI (S), SfiI (F), and BshHII (B) digested YAC 204B2 were hybridized with probes corresponding to the 3' ends of SLC22A1, 2, and 3. The molecular weight marker is indicated at the left.

...ly regulated expression of Slc22a3 indicate a probable role in the metabolic function of the placenta. Towards the end of gestation, a decrease in the levels of Slc22a3 mRNA levels in placenta is observed (Fig. 3B). This may reflect either a temporal limit in gene function or a restriction of Slc22a3 function to a certain cell type for which its transport function remains essential.

The Slc22a3 gene was isolated from the twLub2 deletion region that is associated with two embryonic lethal mutations. Whereas the T-associated maternal effect (Tme) is caused by the absence of the imprinted Igf2r gene (5), the gene mutated in the nonimprinted t^73 allele remains to be identified. Several observations suggest that Slc22a3 presents a good candidate for the t^73 lethal factor. First, the gene maps within the twLub2 deletion, which is defined as the critical region for t^73 (3). Second, Slc22a3 is not imprinted. This was determined by examination of placentas from embryos carrying the T-hairpin deletion (7), which showed that Slc22a3 was expressed equally from both parental alleles (data not shown). Third, in contrast to Slc22a1 and 2, which map to the same genomic locus and are exclusively expressed in adult...
Chapter 2

tissues, Slc22a3 is expressed in placenta, a derivative tissue of the trophoblast cell lineage, that is specifically affected in homozygous r^w73 embryos (6). We are currently testing whether mutations in Slc22a3 are responsible for causing the r^w73 phenotype by the generation of transgenic mice.

In vitro transport assays have shown that rat Slc22a3, as well as the rat and human SLC22A1 and the SLC22A2 proteins, show a broad specificity in the transport of a similar set of organic cation molecules (9,15,17,18). This suggests that the proteins may be functionally redundant, but are expressed through different regulatory mechanisms. Slc22a3, e.g., is unique in possessing an exceptionally long 3' untranslated region, as well as a CpG island at the transcription start site. However, within the Slc22a3 protein sequence, putative N-glycosylation and phosphorylation sites were found, which are not shared among the other Slc22a3 proteins and may contribute to intrinsic differences in transport function (Fig. 2A).

Gene organization of the solute carrier genes

Analysis of a mouse BAC contig spanning 300 kb has revealed that three organic cation transporter genes, Slc22a1, 2, and 3, show a very close physical linkage between the Igf2r and Plg genes in the proximal part of the r^w73 deletion region. This linkage and gene order are also observed for the human gene homologues, and YAC mapping demonstrated that IGF2R, SLC22A1, 2, and 3 and APO(a) lie within a 530-kb region. In mouse, the 3' end of the Slc22a3 gene is separated by approximately 1 kb from the 3' end of the Plg gene. Interestingly, when a BLASTN search was performed with the SLC22A3 EST clones, nine EST clones showed a strong homology over a 206 bp region with terminal exon 19 of the human PLG gene (indicated in Table 1). Each EST clone contains six identical nucleotide changes with respect to the PLG cDNA sequence. Furthermore, the PLG gene, which spans a 53-kb locus (30), maps approximately 250 kb proximal of the APO(a)-like gene (27). Together, these observations suggest that the 3' end of SLC22A3 contains a PLG-like sequence that has duplicated and subsequently diverged. Possibly, this duplication is part of the APO(a)-like gene locus.

Mouse chromosome 17 shares three large homology blocks with human chromosome 6: two on the short arm at 6p21-p21.3 and one on the long arm at 6q21-q27. This latter region, called homology group 67, is syntenic with the region spanning 3.7 - 8.3 cM of mouse chromosome 17 and spans the loci described in this report (see comparative mapping portion of the Mouse Genome Database (http://www.informatics.jax.org) and Debry-Seldin Human/mouse comparative mapping database (http://www.ncbi.nlm.nih.gov/homology/)). The synteny, between mouse and human, of homology group 67 is fully conserved with the exception that the APO(a)/PLG multigene family only contains a single copy of Plg in standard laboratory mouse strains (however, it should be noted that Plg is rearranged to a multicopy family in strains of mice carrying the t haplotype version of chromosome 17; see (24)). The results described here that place three members of the Slc22a gene family into this large homology group highlight the conservation of this region between mouse and human and demonstrate that this multigene family evolved in an ancestor common to both. Homologues of the SLC22A1 and 2 genes show selective high expression in adult liver and kidney, suggesting their expression in a wide range of vertebrates within these organs. In view of the high expression of SLC22A3 in placenta, it will be of interest to test whether the same three-member multigene family exists in nonmammalian vertebrates. The 6q homology 67 group is a chromosomal region that contains the imprinted IGF2R gene plus other loci known to influence coronary artery disease, diabetes, and tumor susceptibility (http://www.ncbi.nlm.nih.gov/
Conserved Slc22a gene cluster in mouse and human

OMIM). The availability of physical maps and gene homologues from this region will help to identify long range cis-acting regulatory elements that may be associated with imprinted gene expression, as well as aid in the future development of mouse models of human disease susceptibility.

**Materials and Methods**

**CpG island mapping**

P1 phage 55D4 was digested with BamHI alone or in the presence of BssHII or SacII to identify BamHI fragments containing CpG-rich sequences. A 4.3-kb BamHI fragment was isolated and partially sequenced (1790 bp; GenBank accession number AF078748). A CpG versus GpC plot revealed a 605-bp CpG-rich sequence between nucleotide positions 681 and 1286.

**Exon trapping**

Exon trapping was performed as described (31). BamHI fragments of P155D4 were randomly subcloned into the exon trap vector pERVF0 and transiently transfected into COS cells. After 40 h, RNA was isolated followed by a reverse transcriptase (RT) reaction, and a nested PCR with primers corresponding to the human β globin exon 2 and exon 3 sequences. PCR products larger than the expected size for a globin exon 2 to exon 3 splicing product were subcloned and analyzed.

**DNA sequencing**

DNA was sequenced using the Dye Terminator Cycle Sequencing kit (Perkin-Elmer) and analyzed with the Perkin-Elmer ABI 373 sequencing device.

**Probes**

The following DNA fragments and primers were used: Slc22a3, 4.3 kb BamHI genomic CpG island subclone and a 740-bp PstI internal fragment, 806-bp cDNA fragment isolated as the most 3' cDNA sequence; Slc22a1, 500-bp 3' and 1760-bp 5' cDNA fragments (8); Slc22a2, primers 5'-agagtcgtctgagctgaccg-3' and 5'-caaccacagcaatacgac-3'; SLC22A1, a 3-kb genomic fragment PCR amplified with primers 5'-ccctcattttgtttgcggtgttgccc-3' and 5'-ctgagttttggacattaagctgtg-3'; SLC22A2, 200-bp cDNA fragment PCR amplified with primers 5'-aacctaggtctacagcagctgg-3' and 5'-gtatttcggtgctagcatcgc-3'; SLC22A3, EST clone ID123234; human APO(a), a 143-bp cDNA fragment PCR amplified with primers 5'-cactgagcagccatgggct-3' and 5'-tttctggtgctatgtgtgg-3' (32); IGF2R, a 9-kb cDNA fragment.

**Southern blotting**

Southern blot hybridizations, including the BAC library screening, were performed under standard Church hybridization and washing conditions (33), using [α-32P]dCTP randomly labeled DNA probes. Oligonucleotides were end-labeled using T4 kinase and [γ-32P]ATP, and hybridized at 38°C.

**Northern blotting**

Total RNA was isolated by the lithium chloride extraction method (34). Fifteen
Chapter 2

micrograms was analyzed on a 1% Mops/formaldehyde agarose gel. After blotting, the membrane was stained with 0.04% methylene blue in 0.5 M sodium acetate, pH 5.2, before it was hybridized under Church hybridization conditions. For SLC22A3, a human RNA MasterBlot was analyzed (Clontech).

Isolation of mouse Slc22a3 cDNA

A degenerate primer (5'-ttgcttgcttg-3', nucleotide positions 1929-1914) was designed against a small region of nearly complete sequence conservation among the different solute carrier 22a family members, and used in a RT-PCR on 12.5 dpc placenta with a forward primer located in the trapped exon (5'--tggcaggctcatcattca-3'). This yielded a predicted size fragment of 1 kb fragment, which was used as a probe to screen a mouse 13.5 dpc placenta cDNA library (obtained from J. Rossant). Five independent clones containing a total cDNA length of 2.7 kb were obtained. All clones ended at the same EcoRI site, suggesting an internal EcoRI site at the 3' end of the cDNA. To obtain the most 3' cDNA sequences, a 3' RACE (rapid amplification of cDNA ends) was performed on 12.5 dpc placenta RT product with two nested Slc22a3 primers (5'-tttgcttgctcgctatgc-3', and 5'-ctgcttttcttcccttgagc-3'). A 1.2-kb fragment was obtained and used as a probe to rescreen the cDNA library. Six independent clones were isolated, all of which contained a 806-bp cDNA insert.

BAC DNA isolation

A gridded mouse BAC library (Research Genetics, Inc) was screened with the Slc22a3 4.3-kb BamHI and trapped exon genomic probes. Five positives were identified with the addresses 27M19 (Bac1), 52B10 (Bac3), 53H6 (Bac4), and 228C21 (Bac9). BAC DNA was isolated by standard alkaline lysis, using 100 ml of solutions I, II and III per 250 ml of bacteria culture. After centrifugation the supernatant was directly applied to a Qiagen tip500 column (Qiagen). On average, a yield of 10-50 ng DNA/ml culture was obtained. For generating a restriction enzyme map, BAC DNA was digested with a combination of rare cutting enzymes followed by pulsed field gel electrophoresis (PFGE; CHEF-DRIII PFGE System, Bio-Rad).

Yeast DNA isolation and analysis

Human YAC clone 204B2 was isolated by R. Taramelli and co-workers (27) and obtained from CEPH (Paris). Yeast DNA preparation and restriction enzyme digestion was done as described (7), followed by PFGE analysis.

GenBank Accession No.: AF078748 (1.79-kb 4.3 BamHI), AF078750 (complete Slc22a3 cDNA ), AF078749 (partial SLC22A3 cDNA).

Acknowledgements

We would like to thank Danielle te Vruchte for excellent technical assistance, Drs Peter Demant and Robert Lyle for reading of the manuscript, and Robert Lyle for help in computer usage, Mehrnaz Ghazvini for providing the mouse placenta Northern blot, and Jeroen de Jong for help in analysis of the mouse BAC contig. We are also indebted to the following persons for providing reagents: Drs S. Varmuza and J. Rossant (mouse 13.5 dpc placenta cDNA library), Dr M. Mannens (human first trimester placenta RNA), Prof. M. Busslinger (BJA-B RNA), Drs P. Valk and R. Delwel (exon trap vector and PCR primers). This work was supported by a grant of the Dutch Cancer Society (KWF) to S.V. and R.Z.
Conserved *Slc22a* gene cluster in mouse and human

**References**

31. Valk, P.J., Hol, S., Vankan, Y., Ihle, J.N., Askew, D., Jenkins, N.A., Gilbert, D.J., Copeland,
Chapter 2
