ORIGINAL ARTICLE

The imprinted region on human chromosome 7q32 extends to the carboxypeptidase A gene cluster: an imprinted candidate for Silver-Russell syndrome

L Bentley, K Nakabayashi, D Monk, C Beechey, J Peters, Z Birjandi, F E Khayat, M Patel, M A Preece, P Stanier, S W Scherer, G E Moore

J Med Genet 2003;**40**:249–256

See end of article for authors' affiliations

Correspondence to: L Bentley, Department of Fetal and Maternal Medicine, Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College, Hammersmith Campus, Du Cane Road, London W12 0NN, UK; l.bentley@ic.ac.uk

Revised version received 5 December 2002
Accepted for publication 6 December 2002

Imprinted gene(s) on human chromosome 7q32-qter have been postulated to be involved in intrauterine growth restriction associated with Silver-Russell syndrome (SRS) as 7-10% of patients have mUPD(7). Three imprinted genes, *MEST*, *MESTIT1*, and *COPG2IT1* on chromosome 7q32, are unlikely to cause SRS since epigenetic and sequence mutation analyses have not shown any changes. One hundred kilobases proximal to *MEST* lies a group of four carboxypeptidase A (*CPA*) genes. Since most imprinted genes are found in clusters, this study focuses on analysing these *CPA*s for imprinting effects based on their proximity to an established imprinted domain. Firstly, a replication timing study across 7q32 showed that an extensive genomic region including the *CPA*s, *MEST*, *MESTIT1*, and *COPG2IT1* replicates asynchronously. Subsequently, SNP analysis by sequencing RT-PCR products of *CPA1*, *CPA2*, *CPA4*, and *CPA5* indicated preferential expression of *CPA4*. Pyrosequencing was used as a quantitative approach, which confirmed predominantly preferential expression of the maternal allele and biallelic expression in brain. *CPA5* expression levels were too low to allow reliable evaluation of allelic expression, while *CPA1* and *CPA2* both showed biallelic expression. *CPA4* was the only gene from this family in which an imprinting effect was shown despite the location of this family of genes next to an imprinted cluster. As *CPA4* has a potential role in cell proliferation and differentiation, two preferentially expressed copies in mUPD patients with SRS syndrome would result in excess expression and could alter the growth profiles of these subjects and give rise to intrauterine growth restriction.

In mammals, a subset of autosomal genes are expressed according to their parental origin. This phenomenon is known as genomic imprinting and is thought to have evolved over at least 100 million years ago.[1] Imprinted gene expression has been shown to be critical for early mammalian embryonic development, growth, and behaviour.

Dysfunctional imprinted gene(s) on human chromosome 7 have been postulated to be involved in the growth restriction associated with Silver Russell syndrome (SRS) as 7-10% of patients have maternal uniparental disomy (mUPD(7)).[2] SRS is a common cause of growth failure in children with an incidence of approximately 1:3000 births, similar to that observed for Turner syndrome.[3] Features include intrauterine growth restriction (IUGR), persistent short stature, characteristic facial features, asymmetry, and fifth finger clinodactyly.[4–6] Most cases of SRS are sporadic with no established Mendelian inheritance pattern.

Most imprinted genes are found in clusters.[7] To date, there are three imprinted regions distributed along the length of human chromosome 7. These are 7p11.3-p13 including growth factor receptor protein 10 (*GRB10*),[8–11] 7q21.3 including the paternally expressed gene 10 (*PEG10*) and epsilon-sarcoglycan (*SGCE*),[12] and an imprinted gene cluster at 7q32 including the mesoderm specific transcript (*MEST*), *MEST* intronic transcript 1 (*MESTIT1*), and non-clathrin coat protein gene intronic transcript 1 (*COPG2IT1*).[13–15]

The human *MEST* gene is imprinted in fetal tissues and localised to chromosome 7q32.2. Two isoforms of the gene are transcribed using alternative first exons. Isoform 1 is paternally expressed in fetal tissues,[13] while isoform 2 is biallelically expressed apart from in the placenta and possibly kidney where it is also paternally expressed.[14 16] The non-clathrin coat protein gene (*COPG2*) is located immediately telomeric to *MEST* in a tail to tail orientation and appears to be biallelically expressed.[15] However, *COPG2IT1* is a transcript located within intron 20 of *COPG2* that is paternally expressed in all fetal tissues examined.[15] An additional untranslated antisense transcript has been identified for *MEST*, namely *MEST* intronic transcript 1 (*MESTIT1*).[14 16] All of these transcripts provide evidence that there is an extensive imprinted cluster at 7q32. Even though most imprinted genes are found in clusters in the genome, this does not exclude biallelic expression of individual genes within such a cluster, such as the testis specific protein A14 (*TSGA14*), which is located directly centromeric to *MEST* and is biallelically expressed in human fetal tissues.[17]

Several lines of evidence have highlighted the region around *MEST* and instigated the search for imprinted genes in this region as candidates for SRS. Not only are 10% of SRS patients mUPD(7), but a single SRS patient has been reported with segmental mUPD(7)q31-qter.[18] The only known imprinted gene within this region that appears to contribute directly to growth is *MEST*. A mouse knock out of *Mest* with a deletion of exons 3 to 9 show pre- and postnatal growth restriction but no asymmetry (one of the commonly found phenotypic characteristics of SRS patients) when paternally inherited.[19] Several groups provided genetic evidence against a role of *MEST* in SRS, reporting the lack of sequence and epigenetic mutations in non-mUPD SRS patients.[20 21] This suggests that other imprinted genes flanking *MEST* at 7q32 may play a role in the pathogenesis of SRS. The connection between imprinting, human 7q, and SRS is strengthened by the fact that the human region is orthologous to mouse proximal chromosome 6. A region in proximal 6, 6A3.2 to 6B3, defined by the breakpoints of the reciprocal translocations T(4;6)77H and T(6;13)6Ad has been shown to be responsible for growth restriction when maternally duplicated.[22]

The transcripts proximal to *MEST* and *TSGA14* belong to the carboxypeptidase A (*CPA*) cluster of genes. The functions of the protein products of the carboxypeptidase gene family range from the digestion of food to the selective biosynthesis of neuroendocrine peptides.[23] Metallo-CPAs are zinc containing exopeptidases, divided into either the N/E or A/B subfamilies based on the overall domain structure and amino acid sequence similarities. All members of the A/B subfamily contain about a 90 amino acid long N-terminal "pro" region that functions as a chaperone to assist with the folding of the active carboxypeptidase (CP) domain. Four members of the A/B subfamily, *CPA1*, *CPA2*, *CPA4*, and *CPA5*, were found to be clustered within an interval of 120 kb on human chromosome 7q32.[24] *CPA1* and *CPA2* (together with *CPB1* on chromosome 3q24) are referred to as pancreatic carboxypeptidases. Although *CPA1* and *CPA2* are expressed most highly in pancreas, expression of the genes in other tissues was also reported.[25] *CPA4* was identified as a gene upregulated in NaBu-treated PC-3 cells (an androgen independent prostate cancer cell line),[26] suggesting the possibility that CPA4 protein may have functions in cell proliferation and differentiation. Although the enzymatic properties of CPA4 have not been reported, the deduced amino acid sequence encoded by the *CPA4* gene shows a high degree of similarity to other CPA proteins. *CPA5* was identified as a new member of the *CPA* gene cluster on 7q32.[24] The gene has been shown to be strongly expressed in testis germ cells. The active site of CPA5 was predicted to cleave substrates with C-terminal hydrophobic residues, as do CPA1 and 2. The four *CPA* genes are only 100 kb proximal to the cluster of imprinted genes (*MEST*, *MESTIT1*, and *COPG2IT1*) on 7q32.[13–16]

As a part of systematic screening for additional imprinted genes in the vicinity of the *MEST*/*COPG2* locus, we have examined the imprinting status of the four *CPA* genes.

MATERIALS AND METHODS

Northern blot analysis

Multiple tissue northern blots containing adult human poly (A)+ RNAs (2 μg/lane) were purchased from Clontech. The cDNA fragments used as probes for *CPA4* and *CPA5* correspond to nt 1440 to 1693 of GenBank acc NM_016352, and nt 1315 to 1420 of NM_080385, respectively. Probes were labelled with [α-32P] dCTP using random priming. Hybridisation and washing were performed following the protocols described by Sambrook and Russell.[27]

Analysis of replication timing

Replication timing was analysed in S phase cells labelled by BrdU incorporation in a normal lymphoblastoid cell line according to the methods described by Hitchins *et al*.[28] Methods used for probe preparation and FISH were based on those described by Harper *et al*[29] and the protocols recommended by Vysis (Vysis, Richmond, UK). Minimums of 200 S phase nuclei were scored for one of three replication patterns: two singlets (single signals representing an unreplicated allele), a singlet plus a doublet (two signals less than two signal widths apart representing a replicated allele), or two doublets. The number of singlet/doublet nuclei were represented as a percentage of the total number of nuclei counted. This reflects the percentage asynchrony of a given genomic region.

Fetal tissue and maternal DNA samples for imprinting analysis

Fetal tissues were acquired following termination of pregnancy at Queen Charlotte's and Chelsea Hospital (QCCH), London. A paired maternal peripheral blood sample was obtained for DNA extraction at the same time. Informed consent was obtained for each sample set. Local ethics approval for obtaining fetal and maternal samples for the study of growth related disorders was granted by the Research Ethics Committee of the Royal Postgraduate Medical School (96/4955). Fetal pancreas tissue was obtained from the MRC Tissue Bank at Hammersmith Hospital. All tissues were washed in sterile PBS, snap frozen in liquid nitrogen, and stored at –70°C. Genomic DNA was extracted from the fetal placenta and maternal blood using the standard phenol-chloroform technique.

PCR amplification and sequence analysis

PCR for genotyping and RT-PCR for expression analysis were performed according to standard protocols in 25 μl reactions supplemented with 1.5 mol/l betaine to increase specificity.[27] PCR and RT-PCR products were purified using microCLEAN (Microzone Ltd). The purified products were sequenced directly using Big Dye terminator kit and resolved on a 3100 Genetic Analyser (Applied Biosystems).

Imprinting analysis in human fetal tissues

Imprinting analysis was carried out by single nucleotide polymorphism (SNP) analysis that requires heterozygous fetal genotypes to differentiate between expression from each of the parental genomes. A homozygous maternal genotype for the same SNP allows identification of the paternal allele since it is the only allele which could have been contributed by the mother. The other fetal allele is therefore paternal in origin. Analysis of fetal RNA by RT, RT-PCR, and DNA sequencing allows us to determine whether both or just one of these alleles is expressed.

CPA1

A C/G polymorphism in the third coding exon (nt173 of BC005279) was found by EST comparison (www.ncbi.nlm.nih.gov/SNP/snpblastByChr.html). Fetal DNA for which matched pancreatic cDNA was available was screened for informative heterozygotes by sequencing genomic PCR products (table 1) with the forward primer used for amplification. RT-PCR products were amplified for informative sample sets. Allele expression was evaluated by sequencing products with the forward RT-PCR primer.

CPA2

A C/T polymorphism for SNP imprinting analysis was identified in exon 7 (nt631 of U19977) by screening the DNA sequence of exons of a set of normal DNA samples. Fetal DNA for which matched pancreatic cDNA was available was screened for heterozygote informativity by sequencing genomic PCR products (table 1) with a nested forward primer (gtgtgaggttgacatgtatc 5′-3′) and reverse PCR primer. RT-PCR products were amplified for informative sample sets. Allele expression was evaluated by sequencing with nested primer caagctacggcactttggac (5′-3′).

CPA4

A G/T polymorphism in exon 9 (nt 914 of AF095719) was identified using the Celera database. Fetal DNA samples, for which matched maternal DNA and a range of fetal tissues were available, were screened for heterozygote informativity by sequencing PCR products amplified from genomic DNA (table 1) using a nested forward primer (caacccttgctccgaagtg 5′-3′). RT-PCR products were amplified for informative sample sets in as many tissue types as were available. Allele expression was evaluated by sequencing with nested forward primer gaggtgaaatcagtggtag (5′-3′) and the reverse, RT-PCR primer.

CPA5

C/T and G/T SNPs were identified in exon 9 (nts1007 and 1014 of AF384667) using the Celera database and used to analyse isoform 1 of *CPA5*. Fetal DNA, for which matched maternal DNA and a range of fetal tissues were available, was screened for heterozygote informativity by sequencing genomic PCR

Table 1 Primer sequences, conditions, and expected product sizes for human genomic and cDNA PCR amplification used for SNP analysis of CP genes

	Genomic DNA PCR:			RT-PCR:		
CPA genes	Primer sequences (5′-3′)	Expected size (bp)	Conditions	Primer sequences (5′-3′)	Expected size (bp)	Conditions
CPA1	F:tgagaacttctggcaccaac R:taccagccatgatgaagagc	401	54°C for 32 cycles	F:aggtacagaaggtgaaggag R:tgtccaaggtgtcgagaatg	540	54°C for 35 cycles
CPA2	F:ttaacattgcaggagaccgg R:cacatctccttaggtccttg	320	54°C for 32 cycles	F:agtgggttacacaagctacg R:agaggcttccagataccttg	198	55°C for 32 cycles
CPA4	F:gagcgtcttgcatttctgtg R:tccctataaagctgcagagc	353	54°C for 32 cycles	F:caacccttgctccgaagtgt R:cataggtcccggtatctctc	359	56°C for 35 cycles
CPA5 Iso 1	F:tggctgccatagtgaacttc R:tttcctgtgaagcggtctc	181	58°C for 32 cycles	F:gcaacagctttgaaaaccag R:ccaaaaatgtactcgatccc	607	60°C for 48 cycles
CPA5 Iso 3	F:tgcaaggttcctaagcctag R:cagggtgtgctccatgatgg	406	60°C for 32 cycles	F:cgggttttggagatgtggcc R:cagggtgtgctccatgatgg	185	64°C for 48 cycles

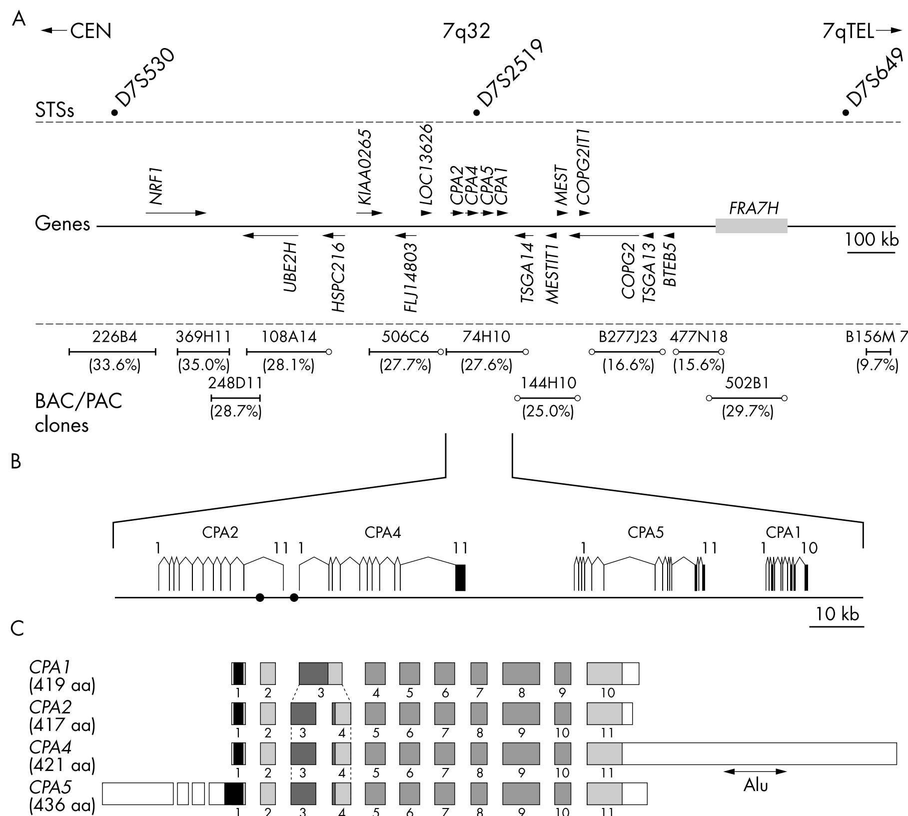

Figure 1 Gene structure of *CPA* genes on human chromosome 7q32, (A) Seventeen known genes/transcripts including three imprinted genes (*MEST*, *MESTIT1*, and *COPG2IT1*) in the 1.5 Mb interval between D7S530 and D7S649. A common fragile site, *FRA7H*,[41] is shown in a hatched box. The arrowheads show the transcriptional orientation of the genes. BAC/PAC clones for replication timing analysis are shown as horizontal bars at the bottom. The percentage asynchrony for each clone is shown in parentheses. B277J23 and B156M7 are BAC clones from the CITB library; the others are PAC clones from the RPCI-1 and -3 libraries.[42] BAC/PAC end positions were determined by end sequences and are represented by vertical bars while those determined by PCR based mapping are shown by open circles. (B) A map for the four CPA genes (*CPA1*, *CPA2*, *CPA4*, and *CPA5*). Spidey (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/), an mRNA to genomic alignment program, was used to determine the exon-intron structures. The mRNA sequences used are NM_001868 (*CPA1*), NM_001869 (*CPA2*), NM_016352 (*CPA4*), and AY155230 (*CPA5*). The genomic DNA sequences used are AC024085 (for *CPA2* and *CPA4*) and AC007938 (for *CPA5* and *CPA1*). *CPA1* consists of 10 coding exons; the other genes consist of 11 coding exons. Two CpG islands (shown by grey circles) were detected by GrailEXP (http://compbio.ornl.gov/grailexp/): one is in intron 10 of *CPA2* (the ratio for observed/expected GC, 0.90; the percent GC, 54.83%); the other is 5′ upstream of exon 1 of *CPA4* (the ratio for observed/expected GC, 0.71; the percent GC, 51.47%). (C) Exons and functional domains in the deduced amino acid sequences encoded by the four CPA genes. 5′ and 3′ UTRs are shown in open boxes. The coding regions are shown in light grey. The signal peptide sequences are shown in black boxes. Carboxypeptide activation peptides and zinc carboxypeptidase domains are shown in dark grey and mid grey boxes, respectively. The repetitive sequence (AluSp) detected by RepeatMasker (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker) in the 3′ UTR of *CPA4* is shown by a bar with arrowheads (nt 1822 to 2110 of NM_016352).

products (table 1) using the nested primer agtgaacttcatcacagccc (5′-3′). RT-PCR products were amplified for informative sample sets in as many tissue types as were available. Allele expression was evaluated by sequencing products with nested forward (agtgaacttcatcacagccc 5′-3′) and reverse (aatgtactcgatccc 5′-3′) primers.

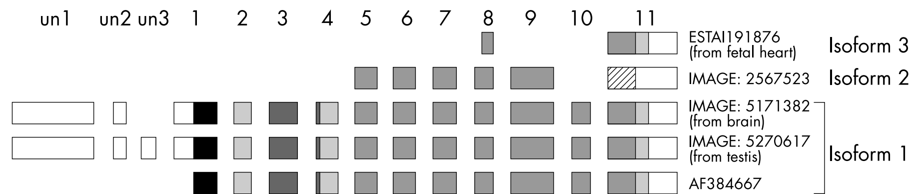

Figure 2 Three splicing isoforms for *CPA5* at its 3′ end. The 11 coding exons for *CPA5* (GenBank acc AF384667) are shown at the bottom. Full length cDNA sequences for IMAGE clones 5171382 (from brain, GenBank acc AY155229) and 5270617 (from testis, GenBank acc AY155230) contain additional three untranslated exons (un1 to 3) besides the 11 coding exons. The mRNA sequences containing all 11 coding exons are referred as "isoform 1". Full length cDNA sequence for IMAGE: 2567523 (GenBank acc AY155228) skips exon 10 (referred to as isoform 2). EST AI191876 (from fetal heart) skips exons 9 and 10 (referred to as isoform 3). UTRs are shown in open boxes. Coding regions are shaded in light grey boxes. The regions corresponding to signal sequence, carboxypeptidase activation peptide (pfam02244), and zinc carboxypeptidase.domain (pfam00246) are shaded in black, dark grey, and mid grey, respectively. The absence of exon 10 in isoform 2 is deduced to cause a frameshift at amino acid 347 and creates 57 amino acids of new sequence showing no similarity with zinc carboxypeptidase.domain (shown in a hatched box).

For isoform 3, a SNP was analysed which lies on exon 11 at nt1132 of AF384667 or nt 78 of the reverse compliment of AI191876. Fetal DNA, for which matched maternal DNA and a range of fetal tissues were available, was screened for heterozygote informativity by sequencing genomic PCR products (table 1) using nested primer aggatgcctctgtaccttgc (5′-3′). RT-PCR products were amplified for informative sample sets in as many tissues as were available. Allele expression was evaluated by sequencing these products with a nested reverse primer (cggaccatcatggagcacac 5′-3′).

Expression analysis of *CPA4* using the pyrosequencing system

Pyrosequencing was used as a quick and accurate method of genotyping fetal and maternal genomic DNA samples and subsequently to quantify allelic expression in RT-PCR products amplified from cDNA generated from fetal tissues of informative subjects. Amplification of fetal and maternal genomic DNA was carried out according to standard PCR protocols using caacccttgctccgaagtgt (5′-3′) and 5′ biotinylated reverse primer tccctataaagctgcagagc (5′-3′) to genotype subjects for the same SNP in exon 9 that was used for analysis by DNA sequencing. The forward pyrosequencing primer aaacatgggaatttcaag (5′-3′) for genomic DNA was designed using SNP Primer Design from Pyrosequencing AB version 1.0.1. For RT-PCR 5′ biotinylated primer caacccttgctccgaagtgt (5′-3′) and reverse primer ctagctggatagacagtggt (5′-3′) flanking exons 10 and 11 were used with a reverse sequencing primer caggtcgatgaagc (5′-3′). Pyrosequencing was carried out according to manufacturers standard protocols (PyroSequencing AB).

RESULTS

Localisation, genomic organisation, and expression of *CPA* genes

The four *CPA* genes, *CPA1*, *CPA2*, *CPA4*, and *CPA5*, are clustered in an interval of 120 kb on 7q32 (*CPA3* is not part of this cluster and has been localised to chromosome 3q24). The distance between the *CPA* genes and a cluster of imprinted genes (*MEST*, *MESTIT1*, and *COPG2IT1*) on chromosome 7 is approximately 100 kb (fig 1A).

Grail EXP was used to highlight potential CpG islands in the human genomic reference sequence AC024083 (RP11-190G13). No CpG islands were localised; however, two small regions were shown to have a slightly higher CpG dinucleotide content than the surrounding DNA (fig 1B, shown by grey circles). The first was located within intron 10 for *CPA2* (102788-103093-PCT GC 54.83) and contained three *Hpa*II (cc^{m}gg) and six *Hin*6I (cg^{m}cg) restriction sites. The second was located 650 bp upstream of the first exon *CPA4* (109418-109648-PCT GC 51.47) and contains only two *Hpa*II restriction sites. Both regions were subjected to RepeatMasker analysis, showing that these regions both align to SINE/Alu elements. The sequence from within intron 10 of *CPA2* consists of two copies of SINE/Alu elements, whereas the second region represents only a single element. Preliminary methylation sensitive PCR showed that both regions are methylated in mUPD(7) and pUPD(7) lymphoblast DNA, as is expected for SINE/Alu elements.[30]

CPA1 has 10 coding exons and the others have 11 coding exons; exon 3 in *CPA1* (234 bp) corresponds to two exons (3 and 4) in the others (fig 1C). The introns 3 of *CPA2*, *CPA4*, and *CPA5* have progressively larger introns containing more repeat elements than one another. *CPA4* has a longer 3′UTR containing a repetitive sequence (Alu) (fig 1C). Interestingly, the 3′UTR of mouse *Cpa4* is also longer and contains repetitive elements (data not shown). Otherwise, the four genes have similar exon-intron structures. All intron-exon boundaries of the four genes conform to the splicing donor/acceptor consensus sequences (GT/AG). Intron phases are completely conserved among the four genes for all exon-intron boundaries. The putative amino acid sequences encoded by the four genes have a similar domain structure: a signal peptide at the N-terminus, carboxypeptidase activation peptide (pfam02244), and zinc carboxypeptidase domain (pfam00246) (fig 1C). *CPA5* has a longer signal peptide than other members (fig 1C[31]). The similarity of amino acid sequences of two of the four proteins ranges from 51% to 68%.

In addition to 11 coding exons of *CPA5*, three non-coding exons at the 5′ end have been reported[24] (un1 to 3, fig 2). By examining ESTs for the *CPA5* gene (Unigene Hs.144699), we found two ESTs indicating the existence of alternative splicing isoforms at the 3′ region of the gene: AW072465 (without exon 10, referred to as isoform 2) and AI191876 (without exons 9 and 10, referred to as isoform 3) (fig 2). Skipping only exon 10 causes a frameshift, and the resulting novel protein sequence after the frameshift did not show significant sequence similarity with any known proteins. Skipping exons 9 and 10 does not shift the reading frame, resulting in a truncated form of the CPA5 protein. By Northern analysis using a multiple human tissue blot (Clontech) with a probe from the 3′UTR, two bands (a major 2.4 kb band and a minor 3.1 kb band) were detected in testis but not other tissues, consistent with the result for mouse *Cpa5* by Wei *et al*[24] (data not shown). The expression of *CPA5* was detected by RT-PCR in many adult and fetal tissues at various levels (data not shown). In RT-PCR analysis using primers common for the three isoforms (a forward primer on exon 8 and a reverse primer on exon 11), isoforms 1 and 3 were detected in many fetal and adult tissues, but isoform 2 was undetectable in the tissues tested.

A probe from 3′UTR of *CPA4* (corresponding to nt 1440 to 1693 of GenBank acc NM_016352) was used to analyse a multiple tissue Northern blot for human adult tissues (Clontech, 7760-1). No signals were detected, a result consistent with that of Huang *et al*.[26] However, using RT-PCR the expression was detected in a wide variety of tissues including placenta, testis, uterus, heart, brain, intestine, kidney, skeletal,

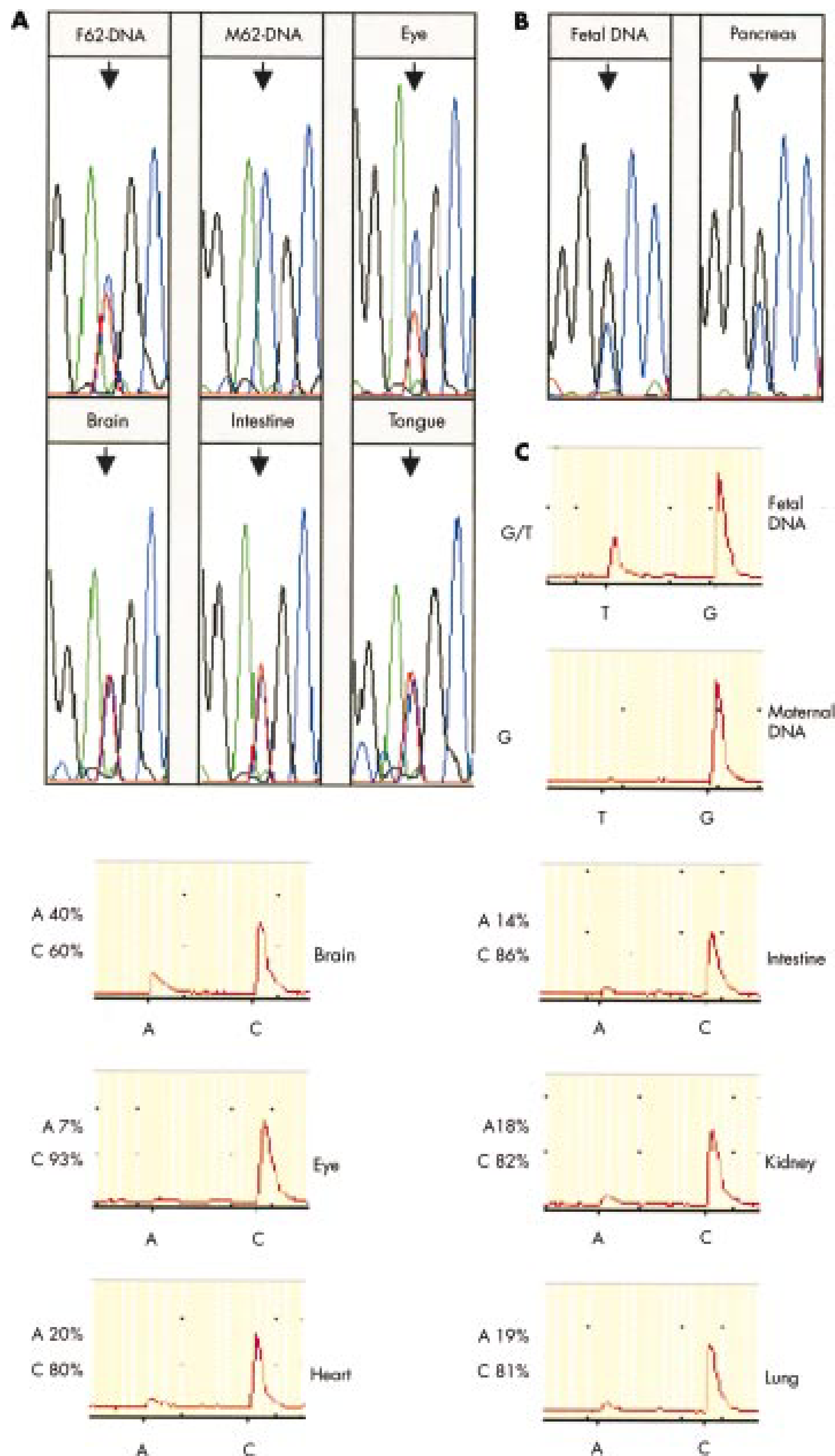

Figure 3 Sequence electropherograms showing SNP analysis of *CPA1*, *CPA2*, and *CPA4*. (A) Sequence electropherograms across the polymorphic site (shown by arrows) in exon 7 of *CPA2* shown for heterozygous fetal and maternal DNA and corresponding cDNA from various tissues as labelled. Expression of both alleles shows biallelic expression for this in all tissues analysed. (B) Sequence electropherograms across the polymorphic site in exon 3 of *CPA1* shown for heterozygous fetal DNA and corresponding cDNA from pancreas tissue showing expression of both parental alleles. (C) Pyrosequencing pyrograms for *CPA4* of fetal and maternal DNA and cDNA of various tissues available for F71. Pyrosequencing of genomic DNA show informativity with a heterozygous fetus and homozygous maternal sample. The G allele is therefore maternal in origin. cDNA alleles are expressed as percentages A and C alleles since a reverse primer was used for this section of the analysis. Taken that a 30% or less contribution of one allele indicates preferential expression, brain is classified as biallelically expressed while eye, heart, intestine, kidney, and lung are preferentially expressed from the maternal allele for this fetus. PCR primers and conditions are shown in table 1.

Table 2 Pyrosequencing results of *CPA4* expressed as percentage contribution of each allele to PCR products analysed. Genotyping of fetal (F) and maternal (M) DNA is expressed as percentages of G and T alleles by sequencing with a forward primer. Results showing informativity for fetal and maternal genotypes are highlighted in bold. cDNA allelic expression is shown as a percentage of C and A alleles since the pyrosequencing primer was in the reverse orientation. For the cDNA expression results show those samples where one allele represents 30% or less of the total expression are in bold

	F34	F46	F47	F50	F58	F65	F66	F67	F71
F DNA	**T 56% G 44% G/T**	**T 58% G 42% G/T**	**T 52% G 48% G/T**	**T 57% G 43% G/T**	**T 58% G 42% G/T**	**T 56% G 44% G/T**	**T 55% G 45% G/T**	**T 54% G 46% G/T**	**T 54% G 46% G/T**
M DNA	T 54% G 46% G/T	T 54% G 46% G/T	T 59% G 41% G/T	**T 0% G 100% G**	**T 0% G 100% G**	T 60% G 40% G/T	**T 0% G 100% G**	T 55% G 45% G/T	**T 0% G 100% G**
Brain		A 58% C 42%	A 39% C 61%					A 45% C 55%	A 40% C 60%
Eye		A 57% C 43%		**A 16% C 84%**	**A 14% C 86%**			**A 86% C 14%**	**A 7% C 93%**
Foot				**A 17% C 83%**					
Heart	A 68% C 32%		A 68% C 32%	**A 19% C 81%**		**A 70% C 30%**			**A 20% C 80%**
Intestine				**A 25% C 75%**		**A 73% C 27%**			**A 14% C 86%**
Kidney	A 61% C 39%				**A 25% C 75%**				**A 18% C 82%**
Lung									**A 19% C 81%**
Placenta		**A 70% C 30%**		**A 4% C 96%**	**A 23% C 77%**		A 48% C 52%	**A 79% C 21%**	
Skin			A 60% C 40%		**A 22% C 78%**		A 45% C 55%		
Spinal C	**A 81% C 19%**								
Tongue			A 68% C 32%			A 67% C 33%			

and muscle (data not shown). *CPA1* and *CPA2* expression was evaluated by RT-PCR in fetal tissues. *CPA1* was expressed exclusively in pancreatic tissue while *CPA2* was ubiquitously expressed.

Imprinting analysis of *CPA* genes in human fetal tissue

Maternal and paternal alleles of imprinted genes replicate at slightly different rates.[32] Generally, genes replicate on the homologous chromosomes in a synchronous manner with 10% of nuclei showing background asynchronous hybridisation patterns. In contrast, imprinted genes showed 25-40% asynchronous replication. This approach can therefore be used to assess imprinting across a large region. In a study by Bonora *et al*,[31] asynchronous replication was shown for genomic clones containing the genes *MEST* and *COPG2.* In our study, replication timing was assessed using 11 PAC and BAC clones across a genomic region of approximately 1.5 Mb extending from *NRF1* to beyond FRA7H. A PAC containing the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene (PAC clone RPCI-1 22F22) was used as the synchronous control where, as expected, only 10% of nuclei showed the background level of asynchronous replication patterns. Each clone analysed representing a region centromeric to and including *MEST* displayed percentage asynchronies of 25% and above (a minimum of 200 nuclei were counted per probe). The PAC clone RPCI-3 502B1, spanning the fragile site FRA7H, showed 29.7% asynchronous replication. Apart from imprinted regions, fragile sites also replicate asynchronously, but not in a parental origin dependent way as is seen in imprinted regions.[33] Replication timing of the BAC clone B277J23 and PAC clone 477N18 from RPCI-3, which lie between the imprinted asynchronous region and the fragile site, most likely corresponds to the telomeric end of the imprinted domain including *MEST*.

Allelic expression of individual *CPA* genes were analysed in fetal tissues by single nucleotide polymorphism (SNP) analysis. Fetal pancreas was the only tissue in which expression of *CPA1* could be detected by RT-PCR. Three out of nine fetal DNA samples were heterozygous for the SNP to be analysed. Sequence analysis showed expression of both parental alleles (fig 3B).

A C/T polymorphism in exon 7 was used to analyse *CPA2* for imprinted expression. Four fetal DNA sample sets which included pancreas were tested and two were found to be heterozygous for the SNP. Both of these showed pancreatic expression of both parental alleles by RT-PCR and DNA sequencing. The same SNP was used to analyse allelic expression in sample sets with a wider range of fetal tissues available. Screening for heterozygotes in these fetal DNA samples showed five informative samples out of 22. Between two and five tissues were available per fetus with a total of seven separate tissues (eye, brain, intestine, lung, placenta, tongue, stomach) between them. Analysis by RT-PCR followed by sequencing showed biallelic expression in all tissues (fig 3A).

For *CPA4* imprinting was assessed for a SNP in exon 9. Preferential expression of one or other parental allele was seen in 75% of the tissue samples analysed (data not shown). Owing to concerns of the sequencing approach not being sufficiently quantitative to measure preferential expression accurately, this analysis was repeated using pyrosequencing as a means of quantifying the molecular contribution of each parental allele to expression in a particular sample (fig 3C). For the purposes of this study preferential expression is defined as a skewed percentage contribution of one allele of 30% versus 70% of the other. Pyrosequencing information was obtained for nine

fetuses in the first trimester with between two and six tissues available per fetus (table 2). Percentage expression for the four brain samples analyses varied from 39/61% to 45/55%, all of which are classified as biallelic according to our definition of preferential expression. Tongue samples were available for F47 and F65, which showed percentage expression of the C allele of 32 and 33% respectively. These, as well as heart samples for F34 and F47, lie within 3 percentage points of the preferential expression cut off. Of the remaining data, 80% of the data may be classified as preferentially expressed according to the definition stated above. It is important to note that since the maternal genotypes were homozygous for the G allele of the SNP, the G allele is maternal in origin for fetuses 50, 58, 66, and 71 (table 2). For F34, F46, F47, F65, and F67, this is not likely to be the case since the T allele is preferentially expressed and therefore according to the pattern of the data, maternal in origin. Each of the allelic expression percentages of tissues, shown in bold, highlights what we believe to be preferential maternal expression.

For *CPA5*, 48 cycles of RT-PCR were required to detect expression in fetal tissues. For each isoform, sequence analysis showed both maternal and parental monoallelic expression and biallelic expression including subjects where each of these patterns was found in different tissues. The data available for each of the isoforms of *CPA5* do not allow a clear pattern of inheritance to be determined. This may be related to the high number of cycles of RT-PCR required for amplification and is not likely to be of biological significance.

DISCUSSION

Four *CPA* genes are localised 100 kb proximal to a cluster of imprinted genes (*MEST*, *MESTIT1*, and *COPG2IT1*) on 7q32.[13–15] Despite *CPA1*, *CPA2*, *CPA4*, and *CPA5* having similar domain structures and exon-intron structures, expression patterns and levels are not uniform between genes. *CPA1*, for example, is only expressed in pancreatic tissue by RT-PCR. Expression levels of *CPA1*, *CPA2*, and *CPA4* are much higher than that of *CPA5*, which requires 48 cycles of RT-PCR amplification for detection. At this level of amplification CPA5 may not be functional in the tissue types analysed.

It had previously been reported that *MEST* replicates asynchronously.[11] In this study, replication timing was used to assess potential imprinting effects across a large genomic region around 7q32 with a series of 11 PAC and BAC clones mapped across the region. In this way it has been possible to define the telomeric boundary of asynchronous replication to between *COPG2IT1* and *TSGA13*. Definition of synchrony versus asynchrony at this boundary is not as clear as expected since it is influenced by asynchronous replication of the fragile site. The entire region as far as the nuclear respiratory factor 1 (*NRF1*) gene, in a centromeric direction relative to *MEST*, replicates asynchronously. Since imprinted genes tend to be found in clusters, the prime candidates for investigation for imprinting effects are the *CPA* genes at 7q32 owing to their proximity to *MEST*.

SNP analysis in human fetal tissues showed biallelic expression of *CPA1* and *CPA2*. *CPA4* showed predominantly preferential expression of the maternally derived allele in a wide range of human fetal tissues except the brain where biallelic expression was found.

Polymorphic imprinting independent of gestational age has been previously reported. For the Wilms' tumour suppressor gene (*WT1*), polymorphic imprinting including biallelic and maternal expression has been seen in placentae while lymphocytes and fibroblasts have displayed polymorphic biallelic and paternal expression.[34–36] Similarly, polymorphic imprinting has been shown for insulin-like growth factor II (*IGF2*) in blood cells[37] as well as for the insulin-like growth factor II receptor (*IGF2R*)[38] and the serotonin-2A (5-HT_{2A}) receptor gene (*HTR2A*).[39] Throughout these reported studies of unusual imprinting patterns, only isolated samples have been described where one allele is partially repressed resulting in preferential expression. In this study, however, *CPA4* SNP analysis has shown a novel pattern of imprinting where the maternally inherited allele is consistently preferentially expressed. Since *CPA4* has been suggested to have a role to play in controlling cell proliferation and differentiation, it may be that a finely controlled balance of expression is required for normal growth. If *CPA4* were to act in limiting growth, then over-expression in a subject with mUPD(7) could have a growth restriction phenotype such as that seen in SRS. This is in agreement with the conflict theory where maternally expressed genes have a growth restricting function and paternally expressed ones act as growth stimulants.[40]

For SNP analysis of *CPA5*, 48 cycles of RT-PCR was required to detect expression of the two transcripts (isoforms 1 and 3) expressed in fetal tissues. This number of cycles of exponential amplification would theoretically result in 1.4×10^{14} amplicons per original copy of template cDNA. According to Sambrook and Russell,[27] 48 PCR cycles can amplify one target molecule to 10 ng of DNA (based on a 200 ng PCR product) when the amplification efficiency is 65-100%. Since in practice more than 10 ng of product would be required, the primer efficiency would have to be somewhat greater than 65%. It is therefore possible that these erratic imprinting results are the result of amplification of a single copy of gene specific cDNA simply because only a single copy was present in the template used for amplification and not as a result of allele specific expression. These results for *CPA5* are therefore not likely to be a true reflection of imprinting expression and are simply an artefact of the limitations of the techniques used.

It is clear that the *CPA* gene family is differentially imprinted and the gene products have different biological functions. It is interesting, despite biallelic expression for *CPA1* and *CPA2*, that *CPA4* is maternally preferentially expressed in a broad spectrum of fetal tissues but not the brain. It is interesting to speculate that a double dose of a gene that is maternally expressed and important in growth and proliferation might cause growth restriction and may be in part accountable for the SRS phenotype. Mutation screening of *CPA4*'s 11 exons is being carried out on DNA from a large cohort of SRS patients.

ACKNOWLEDGEMENTS

The first two authors contributed equally to this work.

Data access. Sequence data from this article have been deposited with the GenBank Data Library under Accession No AY155229 for full length CPA5 cDNA sequence from IMAGE clones 5171382 (brain), AY155230 containing all 11 exons as well as three untranslated exons (isoform 1 from testis, IMAGE: 5270617), and AY155228, a full length cDNA sequence for IMAGE clone 2567523 which skips exon 10 (isoform 2).

REFERENCES

1 **Killian JK**, Nolan CM, Wylie AA, Li T, Vu TH, Hoffman AR, Jirtle RL. Divergent evolution in M6P/IGF2R imprinting from the Jurassic to the Quaternary. *Hum Mol Genet* 2001;**10**:1721-8.
2 **Kotzot D**, Balmer D, Baumer A, Chrzanowsk, K, Hame, BC, Ilyina H, Krajewska-Walasek M, Lurie, IW, Otten BJ, Schoenle E, Tariverdian G, Schinzel A. Maternal uniparental disomy 7 - review and further delineation of the phenotype. *Eur J Pediatr* 2000;**159**:247-56.
3 **Moore GE**, Abu-Amero S, Wakeling E, Hitchins M, Monk D, Stanier P, Preece M. The search for the gene for Silver-Russell syndrome. *Acta Paediatr Suppl* 1999;**433**:42-8.
4 **Silver H**, Kiyasu W, Geirge J, Deamer W. Syndrome of congenital hemihypertrophy, short stature and elevated urinary gonadotrophins. *Pediatrics* 1953;**12**:368-76.
5 **Russell A**. A syndrome of "intrauterine dwarfism" recognised at birth with cranio-facial dysostosis, disproportionally short arms, and other anomalies. *Proc R Soc Med* 1954;**47**:1040-4.
6 **Price S**, Stanhope R, Garet C, Preece M, Trembath. The spectrum of Silver-Russell syndrome: a clinical and molecular genetic study and new diagnostic criteria. *J Med Genet* 1999;**36**:837-42.
7 **Brannan CI**, Bartolomei MS. Mechanisms of genomic imprinting. *Curr Opin Genet Dev* 1999;**9**:164-70.

8 **Jerome CA**, Scherer SW, Tsui LC, Gietz RD, Triggs-Raine B. Assignment of growth factor receptor-bound protein 10 (GRB10) to human chromosome 7p11.2-p12. *Genomics* 1997;**40**:215-16.
9 **Blagitko N**, Mergenthaler S, Schulz U, Wollmann HA, Craigen W, Eggermann T, Ropers HH, Kalscheuer VM. Human GRB10 is imprinted and expressed from the paternal and maternal allele in a highly tissue - and isoform-specific fashion. *Hum Mol Genet* 2000;**9**:1587-95.
10 **Yoshihashi H**, Maeyama K, Kosaki R, Ogata T, Tsukahara M, Goto Y, Hata J, Matsuo N, Smith RJ, Kosaki K. Imprinting of human GRB10 and its mutations in two patients with Russell-Silver syndrome. *Am J Hum Genet* 2000;**67**:476-82.
11 **Hitchins MP**, Monk D, Bell GM, Ali Z, Preece MA, Stanier P, Moore GE. Maternal repression of the human GRB10 gene in the developing central nervous system; evaluation of the role for GRB10 in Silver-Russell syndrome. *Eur J Hum Genet* 2001;**9**:82-90.
12 **Ono**, R, Kobayashi S, Wagatsuma H, Aisaka K, Kohda T, Kaneko-Ishino T, Ishino F. A retrotransposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21. *Genomics* 2001;**73**:232-7.
13 **Kobayashi S**, Kohda T, Miyoshi N, Kuroiwa Y, Aisaka K, Tsutsumi O, Kaneko-Ishino T, Ishino F. Human PEG1/MEST, an imprinted gene on chromosome 7. *Hum Mol Genet* 1997;**6**:781-6.
14 **Nakabayashi K**, Bentley L, Hitchins MP, Mitsuya K, Meguro M, Minagawa S, Bamforth JS, Stanier P, Preece M, Weksberg, R, Oshimura, M, Moore GE, Scherer SW. Identification and characterization of an imprinted antisense RNA (MESTIT1) in the human MEST locus on chromosome 7q32. *Hum Mol Genet* 2002;**11**:1743-56.
15 **Yamasaki K**, Hayashida S, Miura K, Masuzaki H, Ishimaru T, Niikawa N, Kishino T. The novel gene, gamma2-COP (COPG2), in the 7q32 imprinted domain escapes genomic imprinting. *Genomics* 2000;**68**:330-5.
16 **Li T**, Vu TH, Lee KO, Yang Y, Nguyen CV, Bui HQ, Zeng ZL, Nguyen BT, Hu JF, Murphy SK, Jirtle RL, Hoffman AR. An imprinted PEG1/MEST antisense expressed predominantly in human testis and in mature spermatozoa. *J Biol Chem* 2002;**277**:13518-27.
17 **Yamada T**, Kayashima T, Yamasaki K, Ohta T, Yoshiura K, Matsumoto N, Fujimoto S, Niikawa N, Kishino T. The gene TSGA14, adjacent to the imprinted gene MEST, escapes genomic imprinting. *Gene* 2002;**288**:57-63.
18 **Hannula K**, Lipsanen-Nyman M, Kontiokari T, Kere J. A narrow segment of maternal uniparental disomy of chromosome 7q31-qter in Silver-Russell syndrome delimits a candidate gene region. *Am J Hum Genet* 2001;**68**:247-53.
19 **Lefebvre L**, Viville S, Barton SC, Ishino F, Keverne EB, Surani MA. Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest. *Nat Genet* 1998;**20**:163-9.
20 **Riesewijk AM**, Blagitko N, Schinzel AA, Hu L, Schulz U, Hamel BC, Ropers HH, Kalscheuer VM. Evidence against a major role of PEG1/MEST in Silver-Russell syndrome. *Eur J Hum Genet* 1998;**6**:114-20.
21 **Kobayashi S**, Uemura H, Kohda T, Nagai T, Chinen Y, Naritomi K, Kinoshita EI, Ohashi H, Imaizumi K, Tsukahara M, Sugio Y, Tonoki H, Kishino T, Tanaka T, Yamada M, Tsutsumi O, Niikawa N, Kaneko-Ishino T, Ishino F. No evidence of PEG1/MEST gene mutations in Silver-Russell syndrome patients. *Am J Med Genet* 2001;**104**:225-31.
22 **Beechey CV**. Peg1/Mest locates distal to the currently defined imprinting region on mouse proximal chromosome 6 and identifies a new imprinting region affecting growth. *Cytogenet Cell Genet* 2000;**90**:309-14.
23 **Reznik SE**, Fricker LD. Carboxypeptidases from A to Z: implications in embryonic development and Wnt binding. *Cell Mol Life Sci* 2001;**58**:1790-804.
24 **Wei S**, Segura S, Vendrell J, Aviles FX, Lanoue E, Day R, Feng Y, Fricker LD. Identification and characterization of three members of the human metallocarboxypeptidase gene family. *J Biol Chem* 2002;**277**:14954-64.
25 **Normant E**, Gros C, Schwartz JC. Carboxypeptidase A isoforms produced by distinct genes or alternative splicing in brain and other extrapancreatic tissues. *J Biol Chem* 1995;**270**:20543-9.
26 **Huang H**, Reed CP, Zhang JS, Shridhar V, Wang L, Smith DI. Carboxypeptidase A3 (CPA3): a novel gene highly induced by histone deacetylase inhibitors during differentiation of prostate epithelial cancer cells. *Cancer Res* 1999;**59**:2981-98.
27 **Sambrook J**, Russell DW. *Molecular cloning: a laboratory manual.* Cold Spring Harbour, NY: Cold Spring Harbour Laboratory Press, 2001
28 **Hitchins MP**, Bentley L, Monk D, Beechey C, Peters J, Kelsey G, Ishino F, Preece MA, Stanier P, Moore GE. DDC and COBL, flanking the imprinted GRB10 gene on 7p12, are biallelically expressed. *Mamm Genome* (in press).
29 **Harper J**, Delhanty J. FISH in preimplantation diagnosis. In: *Methods in molecular medicine. Molecular diagnosis of genetic disease*. Towowa, NJ: Humana Press Inc, 1994.
30 **Hellmann-Blumberg U**, Hintz MF, Gatewood JM, Schmid CW. Developmental differences in methylation of human Alu repeats. *Mol Cell Biol* 1993, **13**:4523-30.
31 **Bonora E**, Bacchelli E, Levy ER, Blasi F, Marlow A, Monaco AP, Maestrini E. Mutation screening and imprinting analysis of four candidate genes for autism in the 7q32 region. *Mol Psychiatry* 2002;**7**:289-301.
32 **Kitsberg D**, Selig S, Brandeis M, Simon I, Keshet I, Driscoll DJ, Nicholls RD, Cedar H. Allele-specific replication timing of imprinted gene regions. *Nature* 1993;**364**:459-63.
33 **Hellman A**, Rahat A, Scherer SW, Darvasi A, Tsui LC, Kerem B. Replication delay along FRA7H, a common fragile site on human chromosome 7, leads to chromosomal instability. *Mol Cell Biol* 2000;**20**:4420-27.
34 **Jinno Y**, Yun K, Nishiwaki K, Kubota T, Ogawa O, Reeve AE, Niikawa N. Mosaic and polymorphic imprinting of the WT1 gene in humans. *Nat Genet* 1994;**6**:305-9.
35 **Nishiwaki K**, Niikawa N, Ishikawa M. Polymorphic and tissue-specific imprinting of the human Wilms tumor gene, WT1. *Jpn J Hum Genet* 1997;**42**:205-11.
36 **Mitsuya K**, Sui H, Meguro M, Kugoh H, Jinno Y, Niikawa N, Oshimura M. Paternal expression of WT1 in human fibroblasts and lymphocytes. *Hum Mol Genet* 1997;**6**:2243-6.
37 **Giannoukakis N**, Deal C, Paquette J, Kukuvitis A, Polychronakos C. Polymorphic functional imprinting of the human IGF2 gene among individuals, in blood cells, is associated with H19 expression. *Biochem Biophys Res Commun* 1996;**220**:1014-19.
38 **Xu Y**, Goodyer CG, Deal C, Polychronakos C. Functional polymorphism in the parental imprinting of the human IGF2R gene. *Biochem Biophys Res Commun* 1997;**197**:747-54.
39 **Bunzel R**, Blumcke I, Cichon S, Normann S, Schramm J, Propping P, Nothen MM. Polymorphic imprinting of the serotonin-2A (5-HT2A) receptor gene in human adult brain. *Brain Res Mol Brain Res* 1998;**59**:90-2.
40 **Haig D**, Westoby M. Parent-specific gene expression and the triploid endosperm. *Am Natruralist* 1989;**134**:147-55.
41 **Mishmar D**, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui LC, Rosenthal A, Kerem B. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. *Proc Natl Acad Sci USA* 1998;**95**:8141-6.
42 **Hayashida S**, Yamasaki K, Asada Y, Soeda E, Niikawa N, Kishino T. Construction of a physical and transcript map flanking the imprinted MEST/PEG1 region at 7q32. *Genomics* 2000;**66**:221-5.