

γ 2-COP, a novel imprinted gene on chromosome 7q32, defines a new imprinting cluster in the human genome

Nadya Blagitko¹, Ute Schulz¹, Albert A. Schinzel², Hans-Hilger Ropers^{1,3} and Vera M. Kalscheuer^{1,+}

¹Max-Planck-Institut für Molekulare Genetik, Ihnestr. 73, D-14195 Berlin (Dahlem), Germany, ²Institute of Medical Genetics, University of Zürich, Zürich, Switzerland and ³Department of Human Genetics, University Hospital Nijmegen, Nijmegen, The Netherlands

Received June 23, 1999; Revised and Accepted September 14, 1999

DDBJ/EMBL/GenBank accession no. AF157833

We describe a novel imprinted gene, γ 2-COP (nonclathrin coat protein), identified in a search for expressed sequences in human chromosome 7q32 where the paternally expressed *MEST* gene is located. γ 2-COP contains 24 exons and spans >50 kb of genomic DNA. Like *MEST*, γ 2-COP is ubiquitously transcribed in fetal and adult tissues. In fetal tissues, including skeletal muscle, skin, kidney, adrenal, placenta, intestine, lung, chorionic plate and amnion, γ 2-COP is imprinted and expressed from the paternal allele. In contrast to the monoallelic expression observed in these fetal tissues, biallelic expression was evident in fetal brain and liver and in adult peripheral blood. Biallelic expression in blood is supported by the demonstration of γ 2-COP transcripts in lymphoblastoid cell lines with maternal uniparental disomy 7. Absence of paternal γ 2-COP transcripts during embryonic development may contribute to Silver–Russell syndrome. However, on mutation scanning the only γ 2-COP mutation detected was maternally derived. Amino acid comparison of γ 2-COP protein revealed close relation to γ -COP, a subunit of the coatamer complex COPI, suggesting a role of γ 2-COP in cellular vesicle traffic. The existence of distinct coatamer complexes could be the basis for the functional heterogeneity of COPI vesicles in retrograde and anterograde transport and/or in cargo selection. Together, γ 2-COP and *MEST* constitute a novel imprinting cluster in the human genome that may contain other, as yet unknown, imprinted genes.

INTRODUCTION

Genomic imprinting of genes resulting in preferential silencing of the maternal or paternal allele in somatic cells plays a pivotal role in several inherited disorders (1–6) and cancers (reviewed in ref. 7). The epigenetic modifications regulating monoallelic expression and determining imprints are yet to be identified. There is strong evidence that DNA methylation and differences in chromatin structure play a role in maintaining

transcription from only one parental allele (reviewed in ref. 8). For some imprinted genes, somatic regulation has been shown to be controlled by enhancer competition (9) or by antisense transcripts (10–12). The number of imprinted genes that exist in the human genome is as yet unknown; current estimates vary between 100 and 200 (13). To date, only 24 human imprinted genes have been described and a few more in the mouse, hence isolating and characterizing novel imprinted genes is essential for understanding the molecular mechanisms underlying genomic imprinting and their role in human diseases.

Over the past years, it has become apparent that imprinted genes tend to cluster, suggesting coordinated mechanisms controlling gene expression and imprinting. In humans, two major imprinting clusters, each comprising >1 Mb of genomic DNA, have been described to date. One imprinting cluster is on chromosome 11p15, the region involved in Beckwith–Wiedemann syndrome (6), long QT and Jervell and Lange–Nielsen syndromes (14,15). The other cluster, on chromosome 15q11–q13, includes paternally transcribed genes which may play a role in the pathogenesis of Prader–Willi syndrome (5) and *UBE3A*, which is monoallelically transcribed from the maternal allele in the brain and is involved in Angelman syndrome (16–19).

Both hetero- and isodisomy of maternal chromosome 7 (mUPD7) have been described in association with dwarfism and Silver–Russell syndrome (SRS; MIM 180860) (20–22). The finding of maternal heterodisomy implies that an imprinted gene may be involved. In contrast, paternal isodisomy for the entire chromosome 7 has been reported in only two patients. One patient has normal stature and no clinically detectable abnormalities apart from congenital chloride diarrhoea and mild high frequency hearing loss (23). The second patient was growth retarded at 6 months of age and has complete situs inversus and immotile cilia (24). Very recently, a girl with SRS who has partial mUPD7 with biparental inheritance of 7p13–q11 has been described (25). If the putative SRS gene is imprinted, it can probably be ruled out from this chromosomal region.

To date the only imprinted gene identified on human chromosome 7 is *MEST*, which is transcribed from the paternally derived allele in fetal tissues (26,27). Monoallelic transcription is accompanied by parent of origin-specific

⁺To whom correspondence should be addressed. Tel: +49 30 8413 1293; Fax: +49 30 8413 1383; Email: kalscheuer@mpim-berlin-dahlem.mpg.de

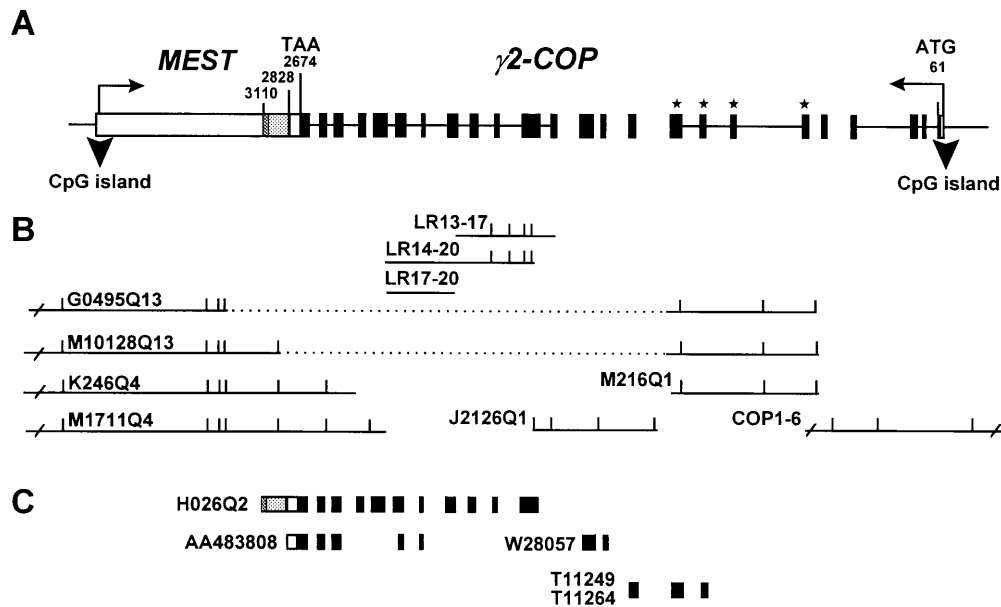


Figure 1. Genomic organization of $\gamma 2$ -*COP* and its location relative to the imprinted *MEST* gene in 7q32. (A) Exon/intron structure of human $\gamma 2$ -*COP*, determined from cosmids (ICRFc113K246Q4, M1711Q4, J2126Q1 and M216Q1), PACs (LLNLP704G0495Q13, M10128Q13 and COP1-6) and by long-range PCR (LR13-17, LR14-20 and LR17-20, numbers referring to $\gamma 2$ -*COP* exons). Black boxes represent translated regions, and open boxes 5'- and 3'-UTRs. The 282 bp longer 3'-UTR found in three $\gamma 2$ -*COP* cDNAs is indicated by grey boxes, of which the dark grey part represents the region that overlaps *MEST* by 52 bp. Exons with asterisks were originally trapped from PAC clone LLNLP704G0495Q13. Introns of known sizes are depicted as horizontal lines between adjacent exons. Transcriptional orientation of $\gamma 2$ -*COP* and *MEST* are indicated by arrows, CpG islands by arrowheads. Positions of the translational start codon (ATG) and stop codon (TAA) are related to $\gamma 2$ -*COP* cDNA (GenBank accession no. AF157833). (B) Contig of long-range PCR products (LR) covering exons 14–17, 14–20 and 17–20, cosmids and PAC clones. A deletion in two PAC clones is indicated by the dotted line. Short vertical lines above each clone represent *Eco*RI restriction sites. (C) $\gamma 2$ -*COP* ESTs that were identified by BLASTN are represented by GenBank accession numbers, clone DKFZp408H026Q2 has been identified by library screening and, like ESTs AI202415 and AI214562 (data not shown), overlaps with *MEST*.

methylation. In the mouse, maternal duplication of proximal chromosome 6, including *Mest*, results in early embryonic lethality (28). More recently it has been shown that paternal transmission of the disrupted *Mest* allele causes embryonic growth retardation and that *Mest*-deficient females show abnormal maternal behaviour (29). Although the function of the *MEST* gene product is still unknown, this finding was particularly interesting in the light of the correlation between mUPD7 and SRS. However, mutation scanning ruled out a major role of *MEST* in SRS (30).

We have asked whether chromosome 7q32 harbours additional imprinted genes. Here we describe a novel imprinted gene, designated $\gamma 2$ -*COP*, that we found to overlap with *MEST*. $\gamma 2$ -*COP* is transcribed from the paternally inherited allele in most fetal tissues, providing clear evidence that imprinting in 7q32 is regional. The precise role of $\gamma 2$ -*COP* protein remains to be elucidated. Its homologue, γ -*COP* codes for a protein that is a component of the heptameric protein complex of COPI vesicles (31). These vesicles are implicated in biosynthetic protein transport within the early secretory pathway between different cellular compartments and are involved in anterograde endoplasmic reticulum (ER) to Golgi transport, intra-Golgi transport, as well as retrograde Golgi to ER transport (reviewed in refs 32–34). Moreover, because $\gamma 2$ -*COP* is a good candidate gene for SRS, we have performed mutation analysis in 42 patients with SRS and in nine patients with primordial growth retardation.

RESULTS

Isolation of $\gamma 2$ -*COP*

We previously reported the isolation and mapping of the imprinted human *MEST* gene. To isolate new genes adjacent to *MEST* we performed exon trapping experiments using PAC clone LLNLP704G0495Q13, which we knew to contain part of *MEST* and had been fluorescent *in situ* hybridization (FISH) mapped to chromosome 7q32 (26). Subsequently, several putative exons were trapped (Fig. 1A). Sequence analysis highlighted exons that originated from *MEST* and four additional exons, two of which were identical to human ESTs (T11264 and T11249 from adult pancreatic islet cells) and two showed similarity to bovine γ -*COP* cDNA (GenBank accession no. X92987) as well as to a group of ESTs and two overlapping contigs (GenBank accession nos THC263483 and THC211632; Institute for Genomic Research, Rockville, MD). Some of the ESTs had been mapped to chromosome 3. Sequence comparisons with bovine γ -*COP* revealed lower per cent identity for our new cDNA (data not shown). Therefore, we designated the newly identified gene $\gamma 2$ -*COP*.

To recover full-length cDNA, we extended our database search by BLASTN analysis using bovine γ -*COP* cDNA as the query sequence for comparison, by cDNA library screening and by partially sequencing *Eco*RI subclones of the genomic region 3' to *MEST*, that were generated from cosmid ICRFc113M1711Q4 (Fig. 1B). Homology searches identified

Table 1. Exon/intron sizes and junctions of the human $\gamma 2$ -COP gene

| Exon no. | cDNA position | Exon size (bp) | Intron size (kb) | 3' splice site | 5' splice site |
|----------|---------------|----------------|------------------|--|-------------------------|
| 1 | 1-97 | 97 | 1.1 | TCGCCG | AGTCTG/ gt acwc |
| 2 | 98-150 | 53 | 0.55 | aaatag/GTAGTG | CAGGAG/ gt atga |
| 3 | 151-231 | 81 | 2.54 | tcatag/GCTAGT | AACCAG/ gt atat |
| 4 | 232-303 | 72 | n.d. | acttag/GGTGAA | AATGAT/ gt aagt |
| 5 | 304-383 | 80 | n.d. | t ttc ag /CAAACA /CAAACA | AAGCAG/ gt atgt |
| 6 | 384-459 | 76 | 3.4 | tcttag/TCTGAC | ACCGAT/ gt aagt |
| 7 | 460-552 | 93 | 1.2 | tattag/GGAACA | TCCCTG/ gt aagt |
| 8 | 553-639 | 87 | 0.9 | at tt ag /CACATG | GTCCAG/ gt attg |
| 9 | 640-797 | 158 | n.d. | ct tt ag /TACCAT | GGATGG/ gt aagt |
| 10 | 798-931 | 134 | <7.0 | tggcag/CCATGA | TTTCAG/ gt ttgt |
| 11 | 932-999 | 68 | <7.0 | tt tt ag /TTCTTC | AACAAG/ gt aggt |
| 12 | 1000-1188 | 189 | n.d. | cacaag/GTGGCA | TTCAAG/ gt aaga |
| 13 | 1189-1284 | 96 | 0.31 | ttacag/GTGGTG /GTGGTG | GATGAT/ gt aagt |
| 14 | 1285-1528 | 244 | 1.8 | tgccag/GGAGGC | GAGCTG/ gt aggt |
| 15 | 1529-1604 | 76 | 0.8 | ttatag/CTGCTG | ACAGAG/ gt aagt |
| 16 | 1605-1708 | 104 | 0.6 | ggacag/GTGTAT | TTAATG/ gt cagt |
| 17 | 1709-1834 | 126 | 1.1 | tcacag/GTTTGA | AAGCAG/ gt tcac |
| 18 | 1835-1897 | 63 | 0.78 | ttatag/AAATCA | TCCAAG/ gt atga |
| 19 | 1898-2037 | 140 | 0.44 | tgatag/AACAAT | TTCCAG/ gt aaga |
| 20 | 2038-2209 | 172 | n.d. | t ttc ag /TTTGAC | CAGCAG/ gt nnnnn |
| 21 | 2210-2307 | 98 | 0.74 | ctccag/TTGCAG | TATGTG/ gt aggt |
| 22 | 2308-2446 | 139 | 0.31 | cactag/CTGGAA | TTGAAG/ gt gaga |
| 23 | 2447-2545 | 99 | 0.47 | acacag/AGGCTG | TGGCAG/ gt aaga |
| 24 | 2546-2676 | 131 | | tctcag/GTATAT | GGATAA |
| 3'-UTR | 2677-3110 | | | | |

The cDNA positions refer to GenBank accession no. AF157833. Bold represents 3' and 5' splice site consensus sequences. n.d., unknown size.

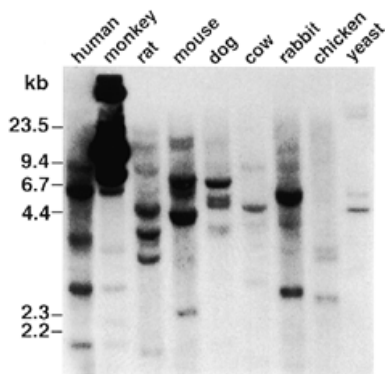


Figure 3. Evolutionary conservation of $\gamma 2$ -COP. Southern blot hybridization of *Eco*RI-digested DNA from various species with $\gamma 2$ -COP cDNA as probe. Strongest signals were observed in monkey, rat, mouse, dog and rabbit. Faint bands are most likely results of cross-hybridization with γ -COP.

2792 (AATTAA and ATTAAA), resulting in shorter transcripts of 2828 bp, and at position 3089 (AATAAA), thereby giving rise to longer transcripts (Fig. 2) (GenBank accession no. AF157833). Attempts to extend the 5'-end did not yield longer products and the position of the putative transcription start site is strongly supported by the similarity with the 5'-end of the bovine γ -COP cDNA.

Evolutionary conservation of $\gamma 2$ -COP was studied by Southern blot hybridization of human $\gamma 2$ -COP cDNA to genomic DNA of other species. The strongest signals were observed with monkey, rat, mouse, dog and rabbit DNA (Fig. 3). Although we had chosen stringent hybridization and washing conditions, yeast DNA that contains only the distantly related γ -COP/*Sec21* gene also yielded weak hybridization signals, which may be due to the much higher number of γ -COP templates in yeast DNA compared with all other DNAs on the blot. Cross-hybridization with human γ -COP could be ruled out by Southern blot hybridization with DNA from suitable human-rodent somatic cell hybrids under the same conditions (data not shown).

In a BLASTX search $\gamma 2$ -COP cDNA disclosed the highest homology to bovine γ -COP and is 71% identical in 2615 bp. γ -COP of *Caenorhabditis elegans* is 60% identical in 2516 bp, of *Arabidopsis thaliana* 61% in 2570 bp and of *Saccharomyces cerevisiae* 56% in 634 bp (data not shown). No significant homology was obtained to γ -COP of *Schizosaccharomyces pombe*.

Genomic organization of human $\gamma 2$ -COP

To establish the complete genomic structure of the human $\gamma 2$ -COP gene, we employed long-distance PCR on genomic DNA and screened a chromosome 7-specific cosmid library and a PAC library. Most exon/intron boundaries were determined by

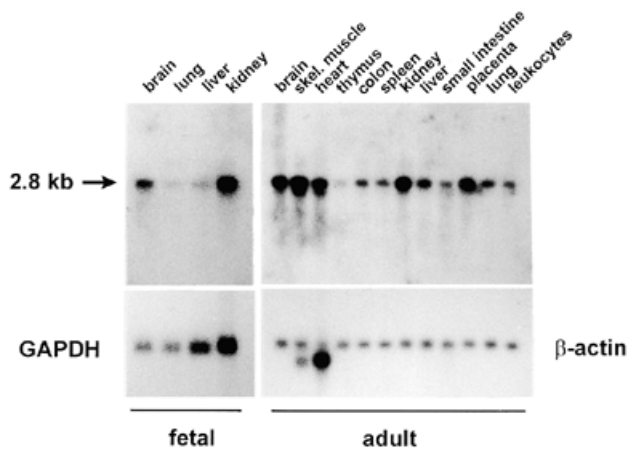


Figure 4. Northern blot analysis of fetal and adult human RNAs. Blots were hybridized to $\gamma 2$ -COP cDNA. A signal of ~ 2.8 kb was present in all tissues examined. Control hybridizations with *GAPDH* (fetal) and β -actin (adult) allowed determination of expression level of $\gamma 2$ -COP.

direct sequencing on genomic clones using oligonucleotides derived from the $\gamma 2$ -COP cDNA. Their selection was greatly facilitated by comparing the human $\gamma 2$ -COP cDNA with the $\gamma 2$ -COP genomic sequence of the pufferfish, *Fugu rubripes* (B. Brunner and V. Kalscheuer, unpublished data). A total of 24 exons were encountered and we found that the complete $\gamma 2$ -COP gene spans >50 kb of genomic DNA. Exon/intron sizes and splice junction sequences are shown in Table 1. Splice sites are in good agreement with the 5' and 3' consensus sequences.

The short 5'-untranslated region (5'-UTR) is G+C rich, has a high frequency of CG dinucleotides and resembles CpG islands. Genomic sequencing upstream of the 5'-UTR region and including part of intron 1 revealed that the CpG island is 630 bp long (data not shown). $\gamma 2$ -COP and *MEST* are orientated in a tail-to-tail configuration and their 3'-UTRs overlap by 52 bp. A 27 bp long repetitive element, resembling L1PA2 repeats, precedes this overlap (Fig. 2). At the same position, a similar element is found in the mouse $\gamma 2$ -COP gene (M. Lang and V. Kalscheuer, unpublished data). To search for conserved synteny, we isolated genomic PAC clones containing the mouse *Mest* gene and found that these clones hybridized with the mouse $\gamma 2$ -COP cDNA, indicating that the gene order has been conserved in mammalian evolution. Detailed characterization of the mouse $\gamma 2$ -COP gene and expression studies are in progress.

$\gamma 2$ -COP protein is a member of a small family

The $\gamma 2$ -COP cDNA contains an open reading frame (ORF) of 2613 nucleotides (nt) beginning with an ATG at nt 61. Comparison of the nucleotide sequence surrounding this putative initiation codon with the Kozak consensus sequence, GCCA/GCCATGG, revealed matches at the conserved positions -5 , -4 and -3 . Translation of the $\gamma 2$ -COP cDNA yielded an ORF of 871 amino acids (Fig. 2). The predicted protein identified five similar proteins in a BLASTP analysis. The highest match with 81% identity was to γ -COP protein of *Bos primigenius* (GenBank accession no. P53620). This alignment supported the proposed start codon, as the two

proteins were in good register throughout their lengths. γ -COP protein of *C.elegans* (GenBank accession no. Q22498) was 61% identical and amino acid sequences of *A.thaliana* (GenBank accession no. O65673), *S.pombe* (GenBank accession no. P87140) and *Saccharomyces cerevisiae* (GenBank accession no. P32074) showed 55, 46 and 38% identity, respectively. A PROSITE search (35) highlighted several presumptive glycosylation sites, one glycosaminoglycan attachment site, one cAMP- and cGMP-dependent protein kinase phosphorylation site, several protein kinase C and casein kinase II phosphorylation sites, one tyrosine kinase phosphorylation site and several *N*-myristoylation sites (data not shown).

Expression of the human $\gamma 2$ -COP gene

In order to determine the expression pattern of the human $\gamma 2$ -COP gene, northern blots containing poly(A) RNA from various adult and fetal tissues were probed with $\gamma 2$ -COP cDNA. We detected a transcript of ~ 2.8 kb in all tissues examined (Fig. 4), suggesting that the established cDNA is full length. Highest expression levels were seen in fetal and adult brain, adult skeletal muscle, heart, kidney and placenta. Fewer $\gamma 2$ -COP transcripts were present in fetal lung, liver and kidney and in adult thymus, colon, spleen, liver, small intestine, lung and peripheral blood leukocytes. To assess the amount of RNA in each lane, the same blots were reprobed with *GAPDH* (fetal) and β -actin (adult).

Imprinting of $\gamma 2$ -COP

To investigate the imprinting status of $\gamma 2$ -COP, we first searched for an intragenic nucleotide sequence polymorphism which would allow paternally and maternally derived alleles in heterozygous PCR samples to be distinguished. Following amplification with sets of primers encompassing the 5'- and 3'-UTR, respectively, single strand conformation polymorphism (SSCP) analysis did not show any bandshift in DNA samples from 10 different individuals. Extending our search with primer sets located in introns, we found a bandshift in intron 22 and sequencing of heterozygous PCR samples revealed a C/A single nucleotide polymorphism. A total of seven fetal DNA samples were heterozygous for this polymorphism, and in four families the parental origin of the transcribed allele could be determined. First strand cDNA was synthesized from the unspliced $\gamma 2$ -COP hnRNA molecules contained in the total RNA fraction from various tissues of all heterozygous fetuses, including skeletal muscle, skin, brain, kidney, adrenal, liver, placenta, intestine, lung, chorionic plate and amnion. Since both primers were located in adjacent introns, mock cDNA synthesis reactions of each RNA sample that included all reagents except reverse transcriptase were performed in parallel and yielded no products. Subsequently, all RT-PCR products were subjected to SSCP analysis. Examples are shown in Figure 5A. In order to prove the concordance between the observed banding patterns and assumed monoallelic or biallelic expression, direct sequencing was carried out in parallel. Examples are shown in Figure 5B. Both comparison of band patterns of genomic DNAs and cDNAs and their sequence analysis indicated monoallelic transcription of the paternal allele in most fetal tissues. Biallelic expression was seen in all fetal brain and liver samples examined,

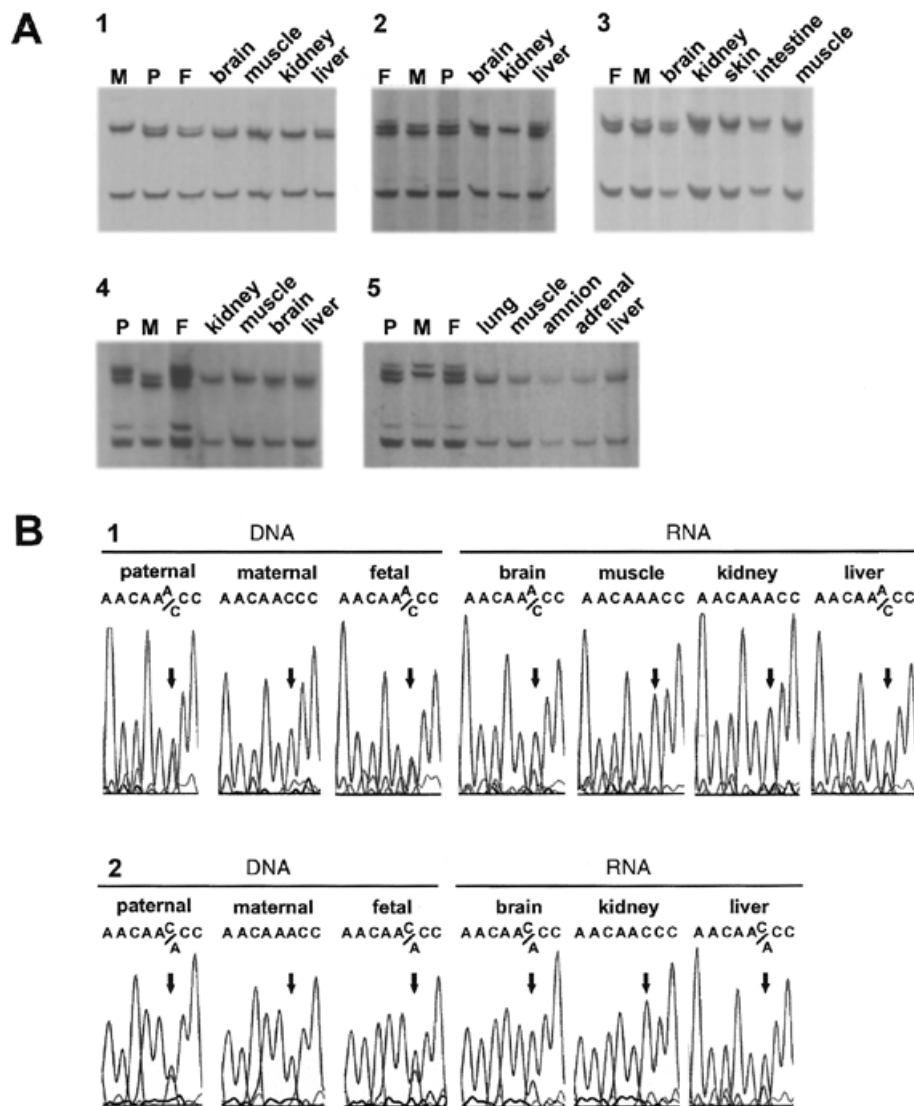


Figure 5. Parental imprinting of $\gamma 2$ -COP in human fetal RNA. Amplification was performed across the single nucleotide C/A polymorphism in intron 22. Products were separated on SSCP gels and partly subjected to direct sequencing. (A) SSCP analysis of five informative fetuses, parental DNA and fetal RNA with opposite genotypic paternal origin of the expressed allele (1, lower allele; 2, upper allele; 3 and 4, upper allele; 5, lower allele). P, paternal DNA; M, maternal DNA; F, fetal DNA. (B) Direct sequencing (samples that are shown in SSCP gels 1 and 2) of opposite genotypic alleles, showing correspondence of SSCP band pattern and allelic transcription of $\gamma 2$ -COP in fetal tissues. The results indicate that $\gamma 2$ -COP transcripts are paternally derived in all fetal tissues investigated, except for fetal brain and liver, which showed biallelic expression.

although barely visible by SSCP in the liver of fetus 5 (Fig. 5A). A summary of all results is given in Table 2. These findings are consistent with genomic imprinting of $\gamma 2$ -COP in fetal tissues, with the exception of brain and liver.

The same polymorphism was used to study allele-specific expression of $\gamma 2$ -COP in adult blood lymphocytes. Interestingly, the two heterozygous samples investigated showed transcripts from both parental alleles. We therefore extended our analysis to total RNA isolated from lymphoblastoid cell lines from two patients with mUPD7. In both cell lines $\gamma 2$ -COP transcripts were present, supporting our previous observation of biallelic expression in blood. In contrast to adult blood lymphocytes, one heterozygous

fibroblast cell line examined showed monoallelic expression of $\gamma 2$ -COP (data not shown).

Mutation scanning in SRS patients

The assignment of the human $\gamma 2$ -COP gene to chromosome 7q32, our finding that it is imprinted and transcribed from the paternally derived allele and its ubiquitous expression prompted us to perform mutation scanning in 42 SRS patients and nine patients with primordial growth retardation. All 24 exons, including exon/intron boundaries, were screened by means of SSCP. Primers and PCR conditions are given in Table 3. Except for five polymorphisms, including two G/A substitutions in the PCR product of exon 17, intron 17 (36/104

DNA samples), a G/A substitution in intron 18 (18/51), a C/T substitution in intron 20 (21/51) and a C/A substitution in intron 22 (26/51), only one bandshift was detected in a patient with SRS (data not shown). Subsequent direct sequencing of the PCR product showed a mutation at position 574 (GAT→CAT) of the $\gamma 2$ -COP cDNA, changing Asp to His. This bandshift has not been detected in other patients nor in 98 DNA samples from healthy controls. SSCP analysis of parental DNA revealed that the mutation was maternally derived, and thus not expressed in most fetal tissues (data not shown).

DISCUSSION

Imprinted genes are not randomly dispersed throughout the genome, but tend to cluster. We have used exon trapping in order to identify genes adjacent to human *MEST*, previously shown by us and others to be imprinted in humans (26,27). This approach led to the isolation of $\gamma 2$ -COP, another imprinted gene. Monoallelic transcription of the paternally inherited allele was evident in all fetal tissues examined, except for brain and liver. Due to limited material, we could not examine $\gamma 2$ -COP expression at earlier stages of human fetal development, which leaves open the question of whether imprinting is not established in brain and liver at all or whether it is lost during early embryonic development. Our finding of a new imprinted gene that overlaps *MEST* defines a new imprinting cluster in the human genome. Both *MEST* and $\gamma 2$ -COP are ubiquitously transcribed from the paternally inherited allele. Their tail-to-tail orientation implies that transcription of *MEST* and $\gamma 2$ -COP is initiated on opposite DNA strands. At present it is unclear whether imprinted transcription of these genes is regulated in a coordinated fashion.

Database searches with human $\gamma 2$ -COP and bovine γ -COP cDNAs revealed related EST sequences that originate from a gene located on human chromosome 3. Sequence alignment of these ESTs resulted in a 2034 bp long cDNA that was more closely related to γ -COP of other species than to human $\gamma 2$ -COP, strongly suggesting that these transcripts represent the bona fide γ -COP gene. The completely sequenced *S.cerevisiae* and *C.elegans* genomes do not contain an orthologue of $\gamma 2$ -COP but the genomes of mouse and pufferfish (*F.rubripes*) do and, as in humans, $\gamma 2$ -COP resides in an orthologous genetic map position relative to *Mest* (M. Lang, B. Brunner and V. Kalscheuer, unpublished data). Based on these findings, it is tempting to speculate that the $\gamma 2$ -COP gene evolved from the ancestral γ -COP form prior to the human-pufferfish divergence or, alternatively, the ancestral gene may have been lost in yeast and nematodes.

In yeast and mammalian cells COPI-coated vesicles mediate Golgi to ER recycling of membrane proteins which contain a dilysine retrieval motif. However, it is clear that COPI also influences transport in the forward direction (reviewed in refs 32–34). The coat of COPI consists of a cytosolic heterooligomeric protein complex, the coatomer, made up of seven subunits, α , β , β' , γ , δ , ϵ and ζ , and of the ADP ribosylation factor (ARF1). It has been shown that it is the γ -COP subunit that binds to both the dilysine motifs of membrane proteins of the early secretory pathway and to the cytoplasmic domain of p23, a major transmembrane protein of COPI vesicles (36–38). It is still not clear how COPI vesicles

Table 2. Overview of fetal tissues examined for allele-specific expression of human $\gamma 2$ -COP, including paternal and maternal genotypes and expression pattern (numbering is identical to that in Fig. 5A)

| DNA/RNA | Tissue | Allele | Expression |
|-----------------|------------------------|--------|-----------------|
| 1. Paternal DNA | | C + A | |
| Maternal DNA | | C + C | |
| Fetal RNA | Brain | C + A | Biallelic |
| | Liver | C + A | Biallelic |
| | Adrenal | A | Paternal |
| | Kidney | A | Paternal |
| | Muscle | A | Paternal |
| | Lung | A | Paternal |
| | Chorionic plate | A | Paternal |
| 2. Paternal DNA | | C + A | |
| Maternal DNA | | A + A | |
| Fetal RNA | Brain | C + A | Biallelic |
| | Liver | C + A | Biallelic |
| | Kidney | C | Paternal |
| | Placenta | C | Paternal |
| 3. Maternal DNA | | C + A | |
| Fetal RNA | Skin | C | Monoallelic |
| | Brain | C + A | Biallelic |
| | Kidney | C | Monoallelic |
| | Intestine | C | Monoallelic |
| | Muscle | C | Monoallelic |
| 4. Paternal DNA | | C + C | |
| Maternal DNA | | A + A | |
| Fetal RNA | Brain | C + A | Biallelic |
| | Liver | C + A | Biallelic |
| | Muscle | C | Paternal |
| | Placenta | C | Paternal |
| | Kidney | C | Paternal |
| 5. Paternal DNA | | C + A | |
| Maternal DNA | | C + C | |
| Fetal RNA | Lung | A | Paternal |
| | Muscle | A | Paternal |
| | Amnion | A | Paternal |
| | Adrenal | A | Paternal |
| | Liver | C + A | Biallelic |
| 6. Fetal RNA | Intestine | A | Monoallelic |
| | Kidney | A | Monoallelic |
| | Muscle | A | Monoallelic |
| 7. Fetal RNA | Kidney | C | Monoallelic |

Fetal RNAs that were informative for determining $\gamma 2$ -COP expression from the paternal allele are shown in bold.

incorporate all of the diverse proteins that reside in the Golgi and how the function of coatomer in both the anterograde and the retrograde direction is regulated. Our finding of a new gene that codes for a γ -COP homologue suggests that in vertebrates at least two distinct coatomer types exist, one containing γ -COP and the other $\gamma 2$ -COP. These distinct coatomers could be the basis for functional diversity in retrograde and anterograde transport and/or in cargo selection, in contrast to the situation in *S.cerevisiae* and *C.elegans*, which have only a single γ -COP

Table 3. PCR primers and conditions used for human $\gamma 2$ -COP mutation scanning

| Exon | Forward primer | Reverse primer | Size (bp) | T_a (°C) | MgCl ₂ (mM) |
|------|----------------------------|---------------------------|-----------|------------|------------------------|
| 1 | ccgagtgaaggcgcggcg | gccggcctgaaagcaggtg | 245 | 66 | 1.5 |
| 2 | gttcactccaagcattcag | attctctgccaccactct | 215 | 60 | 3 |
| 3 | gctataatctgtgatagaaaaggta | aggacacacttatctgagg | 177 | 59 | 3 |
| 4 | tttcagaaaatagaaaatcctgt | gaacactcttagttccctg | 238 | 57 | 1.5 |
| 5 | tttttttgtagatgagactctt | cctattgctttcttatggtc | 308 | 58 | 2 |
| 6 | ttggctaggggtgaactc | aaatcctactaaatctacatggc | 224 | 58 | 3 |
| 7 | ataagagtaattttgcataaacatgt | actcaaaactgtaccatgctta | 206 | 58 | 1.5 |
| 8 | ctctgggaaatgtctgtg | atgctgatcctgagacaag | 229 | 59 | 1.5 |
| 9 | ctttactaatcctgtgctacc | gcaaatggaaaataaccagg | 239 | 59 | 1.5 |
| 10 | ctgtgcttttaacccctatgc | caaaaacttaggaacataaaaagcc | 256 | 60 | 3 |
| 11 | atagttctttaagccaggag | gagagaaaaccaaatgcttcac | 234 | 60 | 3 |
| 12 | ttttttcttgccttccac | ataaataggcagccaacaga | 266 | 57 | 3 |
| 13 | taggaaagtatttttagtagcc | tacttgagagcactactgag | 284 | 58 | 1.5 |
| 14 | aatggcagttttatcttttttg | agcatcgtcaatatcagtgtc | 306 | 58 | 3 |
| 15 | caactgaacatggggacac | cgccctatggtagaatgag | 258 | 61 | 1.5 |
| 16 | tactgtgacatgactattctg | aaaatagctccaagacacc | 199 | 57 | 1.5 |
| 17 | ggcaagagggtataatgctg | agtgacaattaatgagattagag | 257 | 58 | 3 |
| 18 | gctagaaatgtctgctctc | tacaatgcttaggggaaag | 207 | 57 | 1.5 |
| 19 | tccctgtcttctcatagc | tttgtctcaaaagcataccagtt | 242 | 57 | 1.5 |
| 20 | tatgagtaggaagataacttg | ttagaaacctgctgtaggg | 239 | 58 | 3 |
| 21 | cttgcactcccactgat | gaaaaatgttctcaaggca | 220 | 58 | 3 |
| 22 | aaggacttttcttactgactgct | gatcaattagatgatgaagacagg | 330 | 61 | 2 |
| 23 | cacctgtcttcatcatct | tgatggacaacagcaaggc | 378 | 61 | 2 |
| 24 | ccagtgcctgttttgggt | tgcttgatctgctggctc | 333 | 58 | 3 |

Amplification of the first exon was performed including 10% glycerol and 3.5% formamide. All primers are shown 5'→3'. T_a , annealing temperature.

gene. The recent finding of a mammalian ζ -COP-related gene, $\zeta 2$ -COP, the product of which is associated with the COPI complex, and the observation that a given COPI complex contains either ζ -COP or $\zeta 2$ -COP (39) supports the existence of several distinct coatomer forms. Together, these data imply that many or most mammalian cells contain at least two forms of COPI. Location of the $\gamma 2$ -COP protein within the cell is currently being investigated.

Mouse *Mest* maps to proximal chromosome 6 (40). Genetic studies have shown that maternal duplication and paternal deficiency of chromosome 6 proximal to the translocation breakpoint T6Ad cause early embryonic lethality (28). However, knockout experiments have revealed that loss of *Mest* results in viable, but growth-retarded, embryos (29). Based on these results, it is tempting to speculate that lack of transcripts from yet another imprinted gene that maps to proximal chromosome 6 and which is also expressed from the paternally derived allele accounts for fetal lethality. We have found evolutionary conservation of syntenic which may suggest that imprinting is also conserved and that absence of paternal $\gamma 2$ -COP transcripts may be responsible for the imprinting effect observed in the mouse. This contrasts with the situation in humans where mUPD7 is compatible with life. Current investigations of the mouse $\gamma 2$ -COP gene will soon shed more light on this issue.

Ubiquitous transcription of $\gamma 2$ -COP, its imprinting status and chromosomal location prompted us to perform mutation

scanning in SRS patients. However, apart from new polymorphisms we only found one patient who had a mutation in its maternally derived transcriptionally silent allele.

In summary, we have isolated a novel gene, $\gamma 2$ -COP, that is subjected to imprinting in a tissue-specific manner. Ongoing studies in the mouse will soon reveal details of the spatio-temporal expression of this gene and will provide more insight into the involvement of $\gamma 2$ -COP in vesicle traffic. The location of $\gamma 2$ -COP defines a new cluster of imprinted genes in the human genome which may contain other, as yet undetected, genes that are also subject to imprinting. The ongoing search for these genes and their characterization may help to unravel the mechanisms controlling imprinting and its function.

MATERIALS AND METHODS

Identification of human $\gamma 2$ -COP by exon trapping and establishment of the full-length cDNA by library screening and RACE

Following restriction digestion of PAC clone LLNLP704G0495Q1 with *EcoRI*, *PstI* and *SstI*, fragments were shotgun cloned into the pSPL3 vector, provided with the exon trapping system of Gibco BRL (Eggenstein, Germany). All steps were performed according to the supplier's instructions. COS-7 cells (4×10^5) were transfected with 1 μ g of plasmid DNA and total RNA was isolated with Trizol

(Gibco BRL). A total of four $\gamma 2$ -*COP* exons were trapped, identifying two expressed human sequences by BLAST (<http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast>) and the bovine γ -*COP* cDNA by FASTA. Human ESTs containing the 3'-UTR of $\gamma 2$ -*COP* were identified by a BLAST search using the bovine γ -*COP* cDNA as query sequence. A THc database search at the Institute for Genome Research (Rockville, MD) identified two EST contigs, THC263483 and THC211632. These contigs were found to overlap and to correspond to a single cDNA that is homologous to $\gamma 2$ -*COP*.

Human cDNA libraries were obtained from the Resource Center of the German Human Genome Project (Berlin, Germany) and were hybridized with $\gamma 2$ -*COP* cDNA essentially as described previously (26).

5'- and 3'-RACE experiments were performed using the human fetal brain Marathon-Ready cDNA library with the Advantage cDNA PCR kit (Clontech, Heidelberg, Germany). Primers were designed in one of the trapped exons: 241, 5'-AGTAAGCGACTGGCAATTCGGATCAGC-3', and 242 (nested), 5'-TGGCAATTCGGATCAGCATGCAGTAAGC-3', for 5'-RACE; 243, 5'-TCCAGTACCATGCATTGGGAGTC-3', and 211 (nested), 5'-TGCATTGGGAGTGCTGTATCACC-3', for 3'-RACE. Amplifications were carried out as 'touchdown' PCRs with annealing/extension steps of 70→68→66°C for 5'-RACE and 69→67→65°C for 3'-RACE. Products were subcloned into pGEM-T-Easy vector (Promega, Mannheim, Germany) and sequenced in a Li-Cor (MWG-Biotech, Ebersberg, Germany) automated sequencer using standard sequencing primers.

Southern and northern blot hybridizations

For blot hybridizations, DNA probes were labelled by random priming in the presence of [α -³²P]dCTP (Amersham, Little Chalfont, UK). Human multiple tissue northern blots (fetal MTN II, adult 12-lane; Clontech) were hybridized at 42°C in 5× SSPE, 50% deionized formamide, 5× Denhardt's solution, 1% SDS and 10% dextran sulphate with $\gamma 2$ -*COP* cDNA. Washing was in 0.1× SSPE, 0.1% SDS at 65°C and autoradiograms were exposed at -70°C for 16–50 h. Control hybridizations were performed with human β -actin and *GAPDH* cDNA probes.

Southern blot hybridizations were performed at 65°C, a zoo blot (Clontech) was hybridized with $\gamma 2$ -*COP* cDNA probe at 60°C and subsequently washed at the same temperature for 2× 5 min in 2× SSC, 0.1% SDS and for 1× 20 min in 1× SSC, 0.1% SDS.

Polymorphism and allele-specific transcription in fetal tissues by means of SSCP, RT-PCR and direct sequencing

With the informed consent of the parents, biopsies from therapeutic and spontaneous abortions were taken after delivery. DNA was isolated from tissues, blood lymphocytes and cell lines according to standard techniques. PCR was carried out on 100 ng of genomic DNA in a 25 μ l reaction mixture using 10 pmol of each primer, 2 mM MgCl₂, 250 μ M dNTPs and 0.5 U of *Taq* DNA polymerase (Perkin Elmer, Weiterstadt, Germany). The amplification consisted of a total of 35 cycles at 95°C for 1 min, 61°C for 1 min 30 s and 72°C for 2 min. Primers used were 370, 5'-TGATGGGACAACAGCAAGGC-3', and 406, 5'-CACCTGTCTTCATCATCT-3'.

Total RNA was isolated using Trizol (Gibco BRL). First strand cDNA was synthesized from 250–500 ng of total RNA with MMLV reverse transcriptase (Gibco BRL) for 90 min at 37°C using random hexamer primers and was subjected to PCR with a total of 40 cycles under the same conditions as genomic DNA.

For polymorphism and analysis of allele-specific transcription by SSCP, amplification products were loaded onto MDE polyacrylamide gels (FMS BioProducts, Rockland, ME) in 1× TBE, which were run at 500 V for 7 h at 20°C. Silver staining was according to a standard technique. RT-PCR products were sequenced with primers 370 and 406 on an ABI automated sequencer.

Genomic structure of $\gamma 2$ -*COP*

A chromosome 7-specific cosmid library and a human PAC library were screened according to standard protocols. Introns 6–9 and 13–20 were amplified with exon-specific primers using an Expand Long Template PCR System (Roche, Mannheim, Germany) according to the protocol supplied by the manufacturer. PCR amplifications consisted of 35 cycles of 20 s at 92°C, 30 s at 60–65°C (primer-dependent) and 8 min at 68°C; for the last 25 cycles, each extension step was prolonged by 20 s. Exon/intron junctions were determined by direct sequencing with cDNA-specific primers on cosmid and PAC DNA or long-range PCR products, using the *Taq* DyeDeoxy terminator cycle sequencing kit (Applied Biosystems, Weiterstadt, Germany), and analysed on an ABI 377 automated sequencer.

Mutation analysis by means of SSCP

Primers used for mutation scanning of all 24 exons were designed in flanking introns. Primer sequences, sizes of amplification products and PCR conditions are outlined in Table 3. All PCR products were analyzed in 8% polyacrylamide gels containing 2% glycerol using four running conditions: 29:1 and 49:1 acrylamide:bis-acrylamide and 15 and 20°C. Silver staining was performed according to a standard technique.

REFERENCES

- Hall, J.G. (1997) Genomic imprinting—nature and clinical relevance. *Annu. Rev. Med.*, **48**, 35–44.
- Lalande, M. (1996) Parental imprinting and human disease. *Annu. Rev. Genet.*, **30**, 173–195.
- Falls, J.G., Pulford, D.J., Wylie, A.A. and Jirtle, R.L. (1999) Genomic imprinting: implications for human disease. *Am. J. Pathol.*, **154**, 635–647.
- Hurst, L.D. and McVean, G.T. (1997) Growth effects of uniparental disomies and the conflict theory of genomic imprinting. *Trends Genet.*, **13**, 436–443.
- Nicholls, R.D., Saitoh, S. and Horsthemke, B. (1998) Imprinting in Prader-Willi and Angelman syndromes. *Trends Genet.*, **14**, 194–200.
- Reik, W. and Maher, E.R. (1997) Imprinting in clusters: lessons from Beckwith-Wiedemann syndrome. *Trends Genet.*, **13**, 330–334.
- Squire, J. and Weksberg, R. (1996) Genomic imprinting in tumours. *Semin. Cancer Biol.*, **17**, 41–47.
- Tilghman, S.M. (1999) The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell*, **96**, 185–193.
- Leighton, P.A., Saam, J.R., Ingram, R.S., Stewart, C.L. and Tilghman, S.M. (1995) An enhancer deletion affects both H19 and *Igf2* expression. *Genes Dev.*, **9**, 2079–2089.

10. Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F. and Barlow, D.P. (1997) Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature*, **389**, 745–749.
11. Moore, T., Constancia, M., Zubair, M., Bailleul, B., Feil, R., Sasaki, H. and Reik, W. (1997) Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse *Igf2*. *Proc. Natl Acad. Sci. USA*, **94**, 12509–12514.
12. Rougeulle, C., Cardoso, C., Fontes, M., Colleaux, L. and Lalande, M. (1998) An imprinted antisense RNA overlaps *UBE3A* and a second maternally expressed transcript. *Nature Genet.*, **19**, 15–16.
13. Barlow, D.P. (1995) Gametic imprinting in mammals. *Science*, **270**, 1610–1613.
14. Wang, Q., Curran, M.E., Splawski, I., Burn, T.C., Millholland, J.M., VanRaay, T.J., Shen, J., Timothy, K.W., Vincent, G.M., de Jager, T., Schwartz, P.J., Toubin, J.A., Moss, A.J., Atkinson, D.L., Landes, G.M., Connors, T.D. and Keating, M.T. (1996) Positional cloning of a novel potassium channel gene: *KVLQT1* mutations cause cardiac arrhythmias. *Nature Genet.*, **12**, 17–23.
15. Neyroud, N., Tesson, F., Denjoy, I., Leibovici, M., Donger, C., Barhanin, J., Faure, S., Gary, F., Coumel, P., Petit, C., Schwartz, K. and Guicheney, P. (1997) A novel mutation in the potassium channel gene *KVLQT1* causes the Jervell and Lange-Nielsen cardioauditory syndrome. *Nature Genet.*, **15**, 186–189.
16. Vu, T.H. and Hoffman, A.R. (1997) Imprinting of the Angelman-syndrome gene, *UBE3A*, is restricted to brain. *Nature Genet.*, **17**, 12–13.
17. Kishino, T., Lalande, M. and Wagstaff, J. (1997) *UBE3A/E6-AP* mutations cause Angelman syndrome. *Nature Genet.*, **15**, 70–73.
18. Matsuura, T., Sutcliffe, J.S., Fang, P., Galjaard, R.J., Jiang, Y.H., Benton, C.S., Rommens, J.M. and Beaudet, A.L. (1997) *De novo* truncating mutations in *E6-AP* ubiquitin-protein ligase gene (*UBE3A*) in Angelman syndrome. *Nature Genet.*, **15**, 74–77.
19. Malzac, P., Webber, H., Moncla, A., Graham, J.M., Kukulich, M., Williams, C., Pagon, R.A., Ramsdell, L.A., Kishino, T. and Wagstaff, J. (1998) Mutation analysis of *UBE3A* in Angelman-syndrome patients. *Am. J. Hum. Genet.*, **62**, 1353–1360.
20. Kotzot, D., Schmitt, S., Bernasconi, F., Robinson, W.P., Lurie, I.W., Ilyina, H., Mehes, K., Hamel, B.C.J., Otten, B.J., Hergersberg, M., Werder, E., Schoenle, E. and Schinzel, A. (1995) Uniparental disomy 7 in Silver-Russell syndrome and primordial growth retardation. *Hum. Mol. Genet.*, **4**, 583–587.
21. Preece, M.A., Price, S.M., Davies, V., Clough, L., Stanier, P., Trembath, R.C. and Moore, G.E. (1997) Maternal uniparental disomy 7 in Silver-Russell syndrome. *J. Med. Genet.*, **34**, 6–9.
22. Eggermann, T., Wollmann, H.A., Kuner, R., Eggermann, K., Enders, H., Kaiser, P. and Ranke, M.B. (1997) Molecular studies in 37 Silver-Russell-syndrome patients—frequency and etiology of uniparental disomy. *Hum. Genet.*, **100**, 415–419.
23. Höglund, P., Holmberg, C., de la Chapelle, A. and Kere, J. (1994) Paternal isodisomy for chromosome 7 is compatible with normal growth and development in a patient with congenital chloride diarrhea. *Am. J. Hum. Genet.*, **55**, 747–752.
24. Pan, Y., McCaskill, C.D., Thompson, K.H., Hicks, J., Casey, B., Shaffer, L.G. and Craigen, W.J. (1998) Paternal isodisomy of chromosome 7 associated with complete situs inversus and immotile cilia. *Am. J. Hum. Genet.*, **62**, 1551–1555.
25. Miyoshi, O., Kondoh, T., Taneda, H., Otsuka, K., Matsumoto, T. and Niikawa, N. (1999) 47,XX,UPD(7)mat,+r(7)pat/46G,XX,UPD(7)mat mosaicism in a girl with Silver-Russell syndrome (SRS): possible exclusion of the putative SRS gene from a 7p13–q11 region. *J. Med. Genet.*, **36**, 326–329.
26. Riesewijk, A.M., Hu, L., Schulz, U., Tariverdian, G., Hoglund, P., Kere, J., Ropers, H.H. and Kalscheuer, V.M. (1997) Monoallelic expression of human *PEG1/MEST* is paralleled by parent-specific methylation in fetuses. *Genomics*, **42**, 236–244.
27. Kobayashi, S., Kohda, T., Miyoshi, N., Kuroiwa, Y., Aisaka, K., Tsutsumi, O., Kaneko-Ishino, T. and Ishino, F. (1997) Human *PEG1/MEST*, an imprinted gene on chromosome 7. *Hum. Mol. Genet.*, **6**, 781–786.
28. Beechey, C.V. (1999) Imprinted genes and regions in mouse and human. In Ohlsson, R. (ed.), *Genomic Imprinting: An Interdisciplinary Approach. Results and Problems in Cell Differentiation*. Springer Verlag, New York, NY, pp. 303–323.
29. Lefebvre, L., Viville, S., Barton, S.C., Ishino, F., Keverne, E.B. and Surani, M.A. (1998) Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene *Mest*. *Nature Genet.*, **20**, 163–169.
30. Riesewijk, A.M., Blagitko, N., Schinzel, A.A., Hu, L.D., Schulz, U., Hamel, B.C.J., Ropers, H.H. and Kalscheuer, V.M. (1998) Evidence against a major role of *PEG1/MEST* in Silver-Russell-syndrome. *Eur. J. Hum. Genet.*, **6**, 114–120.
31. Stenbeck, G., Schreiner, R., Herrmann, D., Auerbach, S., Lottspeich, F., Rothman, J.E. and Wieland, F.T. (1992) Gamma-COP, a coat subunit of non-clathrin-coated vesicles with homology to Sec21p. *FEBS Lett.*, **314**, 195–198.
32. Gaynor, E.C., Graham, T.R. and Emr, S.D. (1998) COPI in ER/Golgi and intra-Golgi transport: do yeast COPI mutants point the way? *Biochim. Biophys. Acta*, **1404**, 33–51.
33. Lowe, M. and Kreis, T.E. (1998) Regulation of membrane traffic in animal cells by COPI. *Biochim. Biophys. Acta*, **1404**, 53–66.
34. Füllekrug, J. and Nilsson, T. (1998) Protein sorting in the Golgi complex. *Biochim. Biophys. Acta*, **1404**, 77–84.
35. Bairoch, A., Bucher, P. and Hofman, K. (1997) The PROSITE database, its status in 1997. *Nucleic Acids Res.*, **25**, 217–221.
36. Harter, C., Pavel, J., Coccia, F., Draken, E., Wegehngel, S., Tschochner, H. and Wieland, F. (1996) Nonclathrin coat protein gamma, a subunit of coatomer, binds to the cytoplasmic dilysine motif of membrane proteins of the early secretory pathway. *Proc. Natl Acad. Sci. USA*, **93**, 1902–1906.
37. Harter, C. and Wieland, F.T. (1998) A single binding site for dilysine retrieval motifs and p23 within the gamma subunit of coatomer. *Proc. Natl Acad. Sci. USA*, **95**, 11649–11654.
38. Reinhard, C., Harter, C., Bremser, M., Brugger, B., Sohn, K., Helms, J.B. and Wieland, F. (1999) Receptor-induced polymerization of coatomer. *Proc. Natl Acad. Sci. USA*, **96**, 1224–1228.
39. Whitney, J.A. and Kreis, T.E. (1998) ζ 2-COP defines a second form of the COPI complex. *Mol. Cell*, **9**, 339 (abstract).
40. Sado, T., Nakajima, N., Tada, M. and Takagi, N. (1993) A novel mesoderm-specific cDNA isolated from a mouse embryonal carcinoma cell line. *Dev. Growth Differ.*, **35**, 551–560.