

# A novel pleckstrin homology-related gene family defined by *Ipl/Tssc3*, *TDAG51*, and *Tih1*: tissue-specific expression, chromosomal location, and parental imprinting

Dale Frank,<sup>1</sup> Cathy L. Mendelsohn,<sup>2</sup> Emilio Ciccone,<sup>2</sup> Kristian Svensson,<sup>3</sup> Rolf Ohlsson,<sup>3</sup> Benjamin Tycko<sup>1</sup>

<sup>1</sup>Institute for Cancer Genetics and Department of Pathology, Columbia University College of Physicians and Surgeons, 630 W. 168th St., New York, New York 10032, USA

<sup>2</sup>Departments of Urology and Pathology, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA

<sup>3</sup>Department of Genetics and Animal Development, University of Uppsala, Uppsala, Sweden

Received: 10 June 1999 / Accepted: 26 July 1999

**Abstract.** We previously described a gene, *Ipl* (*Tssc3*), that is expressed selectively from the maternal allele in placenta, yolk sac, and fetal liver and that maps within the imprinted domain of mouse distal Chromosome (Chr) 7/human Chr 11p15.5 (Hum Mol Genet 6, 2021, 1997). *Ipl* is similar to *TDAG51*, a gene that is involved in FAS/CD95 expression. Here we describe another gene, *Tih1* (*TDAG/Ipl* homologue 1), with equivalent sequence similarity to *Ipl*. Structural prediction indicates that the products of these three genes share a central motif resembling a pleckstrin-homology (PH) domain, and TIH1 protein has weak sequence similarity to the PH-domain protein SEC7/CYTOHESIN. Like *Ipl*, *Tih1* is a small gene with a single small intron. *Tih1* maps to distal mouse Chr 1 and human Chr 1q31, chromosomal regions that have not shown evidence for imprinting and, in contrast to *Ipl*, *Tih1* is expressed equally from both parental alleles. *Ipl*, *Tih1*, and *TDAG51* have overlapping but distinct patterns of expression. *Tih1* and *TDAG51* are expressed in multiple fetal and adult tissues. In contrast, during early mouse development *Ipl* mRNA and protein are highly specific for two tissues involved in maternal/fetal exchange: visceral endoderm of the yolk sac and labyrinthine trophoblast of the placenta. These findings highlight the dominance of chromosomal context over gene structure in some examples of parental imprinting and extend previous evidence for placenta-specific expression of imprinted genes. The data also define a new subfamily of PH domain genes.

## Introduction

Genomic or parental imprinting is an epigenetic process occurring in the gametes, which leads to a parent-of-origin dependent allelic expression bias of certain genes in the offspring. This phenomenon is documented by an expanding list of imprinted genes in mice and humans, but its mechanism is poorly understood. In particular, the relative importance of gene-specific vs. regional chromosomal features in determining whether a gene is a target for imprinting remains uncertain. There are at least two examples of megabase scale chromosomal “domains” containing multiple imprinted genes: human Chr 11p15.5 and 15q11–q13 and the corresponding areas of distal and middle mouse Chr 7 (see <http://www.mgu.har.mrc.ac.uk/anomaly/anomaly.html> and <http://cancer.otago.ac.nz:810/IGC/web/home.html>). This suggests that chromosomal context may be an important factor in selecting genes for imprinting. On the other hand, transgenic models have suggested that some imprinted genes may contain or be closely

flanked by cis-acting sequences sufficient to mediate imprinting independent of chromosomal position (Chaillet et al. 1995; Pfeifer et al. 1996; Ainscough et al. 1997). Also, from statistical data it was suggested that imprinted genes have, on average, fewer and smaller introns than non-imprinted genes and that this “viral-like” feature may be an important cue targeting these genes for imprinting (Hurst et al. 1996).

Imprinting is conserved in mammalian evolution, so in addition to structural cues biological function may be important in selecting for the presence or absence of imprinting at particular loci. One interesting theory, based on parental conflict, predicts that the set of imprinted genes should be enriched in genes that affect the growth of embryonic and extraembryonic tissues. This theory makes the additional prediction that the direction of imprinting (maternal vs. paternal) should correlate with the direction of the growth effects (positive vs. negative; Haig 1997). While this theory has been challenged (Hurst and McVean 1997), a number of imprinted genes seem to conform to its predictions.

In terms of the predictions of these theories, one interesting imprinted gene is *Ipl* (*Imprinted in placenta and liver*), also known as *Tssc3* (Qian et al. 1997, Lee and Feinberg 1998). This maternally expressed/paternally repressed gene is located in the imprinted domain on distal mouse Chr 7/human Chr 11p15.5. It has an unusually compact gene structure in both mice and humans, consisting in the mouse of two exons and one intron, all contained within 1 kb of DNA (Qian et al. 1997). *Ipl* has sequence homology to *TDAG51*, a mouse cDNA previously shown to be necessary for FAS expression and T-cell receptor activation-mediated apoptosis in lymphocytes (Park et al. 1996; Wang et al. 1998). Here we describe gene structure, imprinting analysis and chromosomal mapping of a second *Ipl* homolog, *Tih1*. We compare the sequences of the three members of this novel gene family and point out the presence of a shared PH domain-like motif. We also survey and compare the tissue-specific expression of these three genes, we examine the cell-type specificity of *Ipl* and *Tih1* mRNA and IPL protein expression in the developing extraembryonic tissues, and we discuss the implications for structural and functional theories of imprinting.

## Materials and methods

**Genomic and cDNA clones.** The cDNA clones corresponding to mouse and human TIH1 ESTs were identified in BLAST searches, obtained from Research Genetics (Huntsville, Ala., <http://www.resgen.com>), and sequenced by standard methods. The genomic mouse *Tih1* clone was obtained by screening a SV/129 genomic  $\lambda$ gt10 library with the full-length *Tih1* cDNA, and the genomic clone was sequenced from multiple internal primers. The genomic and cDNA sequences of mouse *Tih1* and the cDNA

sequence of human TIH1 have been deposited in GenBank (accession nos. AF151099 and AF151100).

**Allelic expression analysis.** Genomic *Tih1* PCR products from BL/6J, spretus, molossinus, and castaneus strains were generated with primers 5'-4 (TAG AGG CTC GGG GGA GGT C) and 3'-1 (GGT GCT CTC TAC GCA CTC C), and these were sequenced to search for nucleotide polymorphisms in the 5' untranslated region of the gene. A single polymorphism, creating a *TaqI* RFLP, was identified at position 523 of the genomic sequence. RT-PCR products were generated with a strategy designed to cross the *Tih1* intron, with primers 5'-4 and 3'-5 (TGG CTG GTT GGC TCC TCC). Genomic DNA and cDNA (Superscript MuLV reverse transcriptase, Gibco-BRL Life Sciences, Gaithersburg, Md.) were prepared from a panel of interspecific hybrid fetal and adult mice as described previously (Qian et al. 1997; Dao et al. 1998). Genomic and RT-PCR products were generated with cycling conditions: 3 min initial denaturation at 94°C; 94°C denaturation 30 s, 56°C annealing 30 s, 72°C extension 1 min, for a total of 38 cycles, with a 7-min final extension step. The gel-isolated PCR products were radiolabeled by an additional eight cycles of PCR and for RFLP analysis were digested with *TaqI* restriction enzyme at 65°C for 3 h, and electrophoresed on 6% non-denaturing polyacrylamide gels. For quantitation of allelic expression by SSCP analysis, the radiolabeled products were digested with *KpnI* and electrophoresed on 8% non-denaturing acrylamide gels at 300 V overnight, at room temperature. Allele-specific expression was quantitated by phosphorimaging (Storm Phosphorimaging System, Molecular Dynamics, Sunnyvale, Calif.).

**Chromosomal mapping.** The chromosomal localization of *Tih1* was determined by typing a panel of genomic DNAs from 94 F<sub>1</sub> generation mice from an interspecific backcross (C57BL/6J × SPRET/Ei) × SPRET/Ei created at The Jackson Laboratory (BSS panel 2; Rowe et al. 1994). The genotyping was done with genomic *Tih1* PCR, as above, followed by restriction digestion to score the *TaqI* RFLP.

**Northern analysis.** Total RNA from fetal and adult mouse tissues was prepared with Trizol™ reagent (Gibco-BRL Life Sciences). For the fractionation of placenta into maternal and fetal sides, 14.5-dpc placentas were snap-frozen in polyethylene glycol-based embedding medium (Stephens Scientific, Riverdale, N.J.). These were oriented with the fetal side down, presenting the maternal layer as a dome-shaped covering over the fetal layers. Sequential 1-mm sections of tissue were taken, starting from the maternal surface and progressing deeper, into the fetal portion of the placenta. In the deeper sections, the thin outer rim of maternal tissue was dissected off and discarded. These sections were lysed directly in Trizol™ and total RNA was prepared. Parallel sections were taken and embedded for standard histological sectioning in transverse and horizontal planes. By visual estimation, the first section was at least 90% maternal tissues, while the third section was entirely fetal tissues. For Northern blotting, 6–10 µg of total RNA per lane was electrophoresed on 1% agarose gels containing formaldehyde and transferred to nylon membranes. Poly-A+ RNA blots were purchased from Clontech (Palo Alto, Calif.). Mouse and human cDNA probes were labeled with α-<sup>32</sup>P (Random Primers DNA Labeling System, Gibco-BRL), and the blots were hybridized in Hybrisol I (Inter-gen, Purchase, N.Y.) at 42.5°C overnight.

**RNA in situ hybridization.** Whole mouse conceptuses were fixed in paraformaldehyde, histological sections were prepared, and the slides were subjected to isotopic (<sup>35</sup>S labeling) or non-isotopic (digoxigenin labeling) in situ hybridization, as described (Ohlsson et al. 1994; Mendelsohn et al. 1999). Anti-sense and sense riboprobes were synthesized from the full-length *Ipl* and *Tih1* cDNA clones and from an Image Consortium cDNA clone corresponding to the 4311 gene. The sense control probes did not give significant signals.

**Western blotting and immunohistochemistry.** The anti-*Ipl* antibody was prepared by immunizing rabbits with a synthetic peptide (CQDF-PRYRYQRSESEMPSE) corresponding to the C-terminal region of the predicted protein (Research Genetics). The anti-*Tih1* antibody was prepared by immunizing with the peptide (CQTVRARQSLGTGLVS). The amino-terminal cysteines were artificial and were included for conjugation to KLH. The antibodies were verified as specific by ascertaining that

specific bands on the Western blots were reduced in intensity when the antiserum was pre-absorbed with excess peptide. Western blotting results were further verified by ascertaining that the pre-immune sera did not yield background bands at the cognate positions on the blots. The Western blots were developed with Super Signal chemiluminescent reagent (Pierce, Rockford, Ill.). Immunohistochemical staining was done on deparaffinized formaldehyde-fixed sections of whole mouse conceptuses after antigen retrieval by microwaving. Antigen detection was by the biotin/avidin-HRP method.

## Results

*Tih1*, a homolog of *Ipl* and *TDAG51*, is a small gene with a single intron. A BLAST search of the EST database (dbEST) using the *Ipl* cDNA sequence identified a cluster of ESTs corresponding to a novel murine gene, which we refer to as *Tih1* (*TDAG/Ipl* homolog-1). A 1419-nt cDNA sequence was determined from the longest of the corresponding cDNA clones and confirmed by sequencing a series of RT-PCR products generated with *Tih1* primers from BL/6 mouse tissues. This yielded an open reading frame encoding a predicted 125-amino acid protein. A presumably full-length 1.6-kb human clone was also identified in this screen and subsequently sequenced. This revealed an open reading frame of 127 amino acids, with complete amino acid sequence identity over all but two residues (an insertion of threonine-alanine in the human sequence after amino acid position 5). Subsequent Northern and Western blotting indicated that the sizes of the mRNA transcripts and proteins were consistent with the size and sequences of the cDNA clones (see below).

To further confirm the murine cDNA sequence and to characterize the exon/intron structure of the *Tih1* genomic locus, we used the *Tih1* cDNA probe to isolate a genomic clone from an Sv/129 λgt10 library. Sequencing of this clone, as well as long PCR products generated from BL/6 genomic DNA, showed that *Tih1* consists of two exons of 770 bp and 776 bp, separated by a single intron of 1580 bp (Fig. 1a). The structure of the human locus was assessed by genomic PCR, which revealed a similarly sized intron of approximately 2 kb. The entire coding sequence in both the human and mouse genes was found to lie within the first exon (Fig. 1a and data not shown).

Two additional noteworthy features in both the human and mouse genes are a translation initiation codon in a non-canonical context (CCCATGA) and a non-canonical polyadenylation signal sequence (AATGAA) (Fig. 1a and data not shown). Both of these variant sequences can be functional in other genes (Kozak 1996; Wu et al. 1993). For comparison, the human IPL gene also utilizes a START codon in a non-canonical context (GACATGA), while the murine version of this gene utilizes a canonical START (Qian et al. 1997). The *TDAG51* gene also utilizes a non-canonical START (AGGATGC), and has an alternative polyadenylation signal sequence (AATTAA; Park et al. 1996). Arguing for the legitimacy of the *Tih1* START codon, an antipeptide antibody detected a protein of the predicted molecular weight on Western blots of murine tissues (see below).

The legitimacy of the unusual polyadenylation signal in *Tih1* was supported by a BLAST search, which showed that at least four additional mouse ESTs corresponding to this gene (for example, accession nos. AI462216, W65066) had their poly-A tracts at the same position 20 to 25 bp downstream from this motif. This variant sequence was also present in five independent rat *Tih1* ESTs (for example, accession nos. AI411461, AI37889) and in nine human TIH1 ESTs (for example, accession nos. AA677613, AA576539) at an appropriate distance within 25 bp upstream of the poly-A tracts. Sequencing of the mouse genomic *Tih1* clone showed no evidence of other (canonical) poly-A addition signals within 250 bp downstream of the AATGAA motif.



*Tih1*, *Ipl* and *TDAG51* define a novel gene family with a shared PH domain-like motif. Amino acid sequence alignments showed that similarity between IPL (144 amino acid residues), TDAG51 (261 aa), and TIH1 (125 aa) is limited to a central stretch of approximately 95–100 aa (Fig. 1b, d). Of the three proteins, IPL and TIH1 have nearly identical structures, consisting of the central conserved domain, which is 53% identical between the two proteins, preceded and followed by short divergent N- and C-terminal sequences. In contrast, TDAG51 and its human homolog (called PQ-rich protein) are larger proteins whose central domains are interrupted by repetitive proline/glutamine-rich sequences and whose carboxy termini are composed of long proline/glutamine/histidine-rich tracts not found in the other two family members (Fig. 1b, d). If the repetitive motifs are removed from the TDAG51 sequence, then the central domain of this protein is 52% and 60% identical to IPL and TIH1, respectively. Screening of the EST database using nucleotide and amino acid sequences from these genes identified numerous human, mouse, and rat ESTs, all of which matched one of the three genes described and failed to reveal any other members in this family. No ESTs or genomic sequences from more divergent organisms were identified, suggesting that the *Ipl/Tih1/TDAG51* gene family is a recent evolutionary innovation.

Prediction of tertiary structure, using the 123D algorithm (<http://genomic.sanger.ac.uk/123D/run123D.shtml>), weighted in favor of sequence and secondary structure, showed that the central domains of *Ipl* and *Tih1* scored most highly versus the PH domain prototype structure pleckstrin (pls) (Z scores of 3.5 and 4.8, respectively). Also identified with significant scores was the PH domain of dynamin (dynA). The central domain of TDAG51, which differs structurally from those of *Ipl* and *Tih1* by an insertion of a proline/glutamine-rich stretch of 35 amino acids, also identified pleckstrin and dynamin, although less significantly (Z score for pleckstrin = 2.0). Although the 123D algorithm ranked other structures, only the PH domain structures were consistently identified with high scores in searches with all three family members. As a point of comparison, the bona-fide PH domain of human SEC7/cytohesin, which has weak homology to *Tih1* (see below), identified pleckstrin with a Z score of 7.0. The SMART algorithm (<http://coot.embl-heidelberg.de/SMART/>), which compares amino acid sequences to a database of various protein motifs, consistently identified PH domains in all three family members. An alignment of the three family members with the PH domains of pleckstrin and dynamin (<http://www.embl-heidelberg.de/~wilmanns/structures/ph.html>) is shown in Fig. 1b.

The PH domain is defined by a canonical secondary structure (seven beta strands followed by a single alpha helix) and a characteristic tertiary structure; high amino acid sequence identity is not expected in alignments of different PH domains (Shaw 1996; Lemmon and Ferguson 1998). This is illustrated by considering proteins that contain the related phosphotyrosine binding (PTB) domain. PTB domains, which are found in adapter molecules like Shc and IRS-1, have little sequence similarity to each other or to "typical" PH domains, but by structural studies they are essentially identical to PH domains (Lemmon and Ferguson 1998; Eck et al. 1996). Nonetheless, there are several conserved amino acids among bona fide PH domains. The most highly conserved of these, a glycine in the most N-terminal beta strand and a tryptophan in the C-terminal alpha helical region, are present in all three family members (Fig. 1b).

PH domains are common among eukaryotic proteins and mediate associations with biological membranes, via binding to phosphatidyl inositol lipids (Shaw 1996; Lemmon and Ferguson 1998). PH domain-containing proteins include examples that are involved in diverse processes including signaling from the plasma membrane, membrane-cytoskeletal interactions, and membrane transport between organelles. Indeed, further evidence for a PH domain-like structure was found in a BLASTp search with the TIH1

protein sequence, which uncovered a short stretch of high similarity (62% similarity, 43% identity) to the N-terminal half of the PH domain of mouse, rat, and human SEC7/cytohesin proteins (Fig. 1c). These proteins promote guanine nucleotide exchange by small GTP binding proteins during endocytic and Golgi membrane flow (Schimmoller et al. 1997; Chardin et al. 1996; Franco et al. 1998; Meacci et al. 1997). Cytohesin also functions at the cytoplasmic face of the plasma membrane to modulate the interaction of the leukocyte integrin  $\alpha$ LB2 with its ligand, ICAM-1 (Kolanus et al. 1996; Nagel et al. 1998).

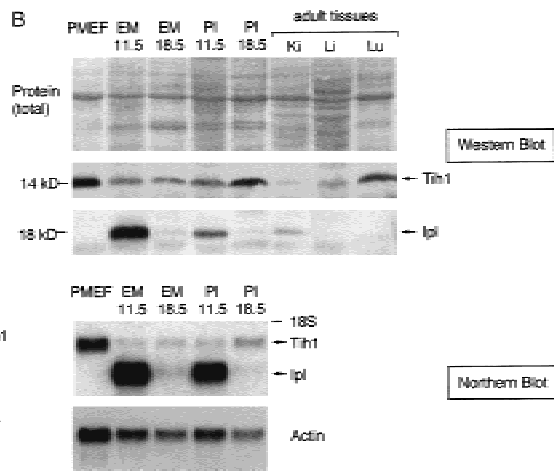
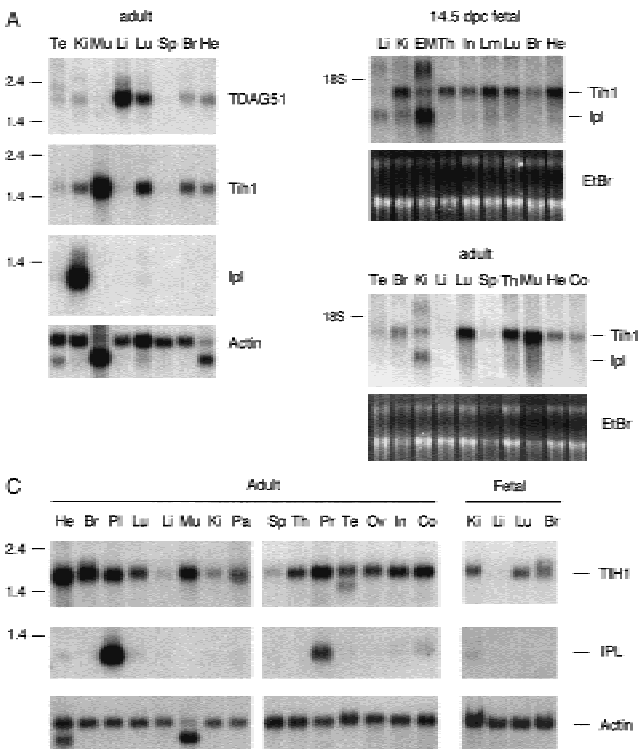
*Tih1* is expressed biallelically in embryonic, extraembryonic, and adult tissues. *Ipl* is a small gene with a single intron, and it is strongly imprinted, with expression predominantly from the maternal allele, in early to mid-embryonic and extraembryonic tissues (Qian et al. 1997). Since one model for genomic imprinting invokes exon/intron structure as influencing the probability of imprinting, the structural similarity of *Tih1* to *Ipl* raised the question of whether *Tih1* might also be imprinted. To test this, we used allelic expression analysis in interspecific mouse crosses, an approach similar to that used previously to show imprinting of *Ipl* and *Impt1* (Qian et al. 1997; Dao et al. 1998). Screening of genomic PCR products by SSCP analysis and sequencing revealed an A  $\rightarrow$  G polymorphism at nucleotide position 523 of the *Tih1* genomic sequence. This change deletes a *TaqI* restriction site, creating an RFLP that can be used to distinguish the *Tih1* alleles of BL/6 (*Mus mus domesticus*) mice from those of non-domesticus strains (*M. m. spretus*, *M. m. castaneus*, and *M. m. mollosinus*) (Fig. 2a). This allowed us to study expression of maternal vs. paternal alleles of *Tih1* in offspring from crosses between these strains. The results for multiple crosses of BL/6 with either CAST or MOLD indicated that expression of *Tih1* is biallelic in multiple tissues at all stages of development examined, including 9.5 dpc, 12.5 dpc, and adult (Fig. 2b and data not shown). These results were verified by an independent method of allelic discrimination, SSCP analysis (Fig. 2b and data not shown). Quantitation by phosphorimaging, with normalization to the results from PCR of heterozygous genomic DNA, showed that for every organ and tissue except placenta the signals from the two parental alleles were equal to within 10 percent. In contrast, using the same crosses, but assessing allelic representation in *Ipl* mRNA, many of these tissues were previously shown to have a pronounced and reproducible allelic expression bias (Qian et al. 1997; Dao et al. 1999).

The only tissue that showed an allelic bias for *Tih1* expression, dependent on the direction of the crosses, was placenta (Fig. 2b, c). We have several reasons to believe that the bias in this tissue reflects contamination of the late-stage placentas with maternal decidual cells, and overrepresentation of the maternal allele owing to production of *Tih1* mRNA from these maternal cells. First, a significant portion of the late gestation placenta (after 11.5 dpc) consists of maternal decidua and blood vessels. This is confirmed by the allelic representation in genomic DNA extracted from these placentas, which shows a consistent bias in favor of the maternal allele (data not shown). Second, a maternally enriched tissue fraction from 14.5-dpc mouse placentas expresses nearly as high a level of *Tih1* mRNA as does a fetally enriched fraction; in striking contrast, there is essentially no detectable message for *Ipl* in the same maternal fraction (Fig. 2c). Lastly, there is no evidence of allelic bias in *Tih1* expression in message from whole 9.5 dpc conceptuses, a developmental stage at which contamination of the dissections with maternal tissue is minimal (Fig. 2b).

*Tih1* maps to murine Chr 1 and human Chr 1q31. We next determined the chromosomal position of *Tih1*, using a backcross panel, (C57BL/6J  $\times$  SPRET/Ei)  $\times$  SPRET/Ei, The Jackson Laboratory, BSS panel 2; Rowe et al. 1994), and taking advantage of the same nucleotide polymorphism (the A  $\rightarrow$  G at nt 523 of the *Tih1*







**Fig. 4.** Tissue-specific expression of the three family members at the mRNA and protein levels. **A**, Expression of *TDAG51*, *Tih1*, and *Ipl* mRNAs in mouse tissues. The blot of adult tissues contains poly-A+ RNA and was hybridized sequentially with the indicated cDNA probes. The blots of fetal tissues contain total RNA and were hybridized with combined *Tih1* and *Ipl* cDNA probes. Of the three genes, *Ipl* has the most restricted tissue distribution. **B**, Comparison of results by Northern and Western blotting. Total protein on the membrane was visualized by Ponceau Red staining. IPL protein is specific for early extraembryonic tissues, while TIH1 protein has a wider tissue distribution. Parallel Northern blot analysis of total RNA shows a direct correlation between the mRNA and protein levels. PMEFL = primary murine embryo fibroblasts; other abbreviations as in Fig. 2. **C**, Expression of IPL and TIH1 mRNAs in human tissues. The blots contain poly-A+ RNA and were hybridized sequentially with the indicated probes. Te = testis, Co = colon, Pa =

pancreas, Pr = prostate, Ov = ovary, In = intestine; other abbreviations as in Fig. 2.

extraembryonic tissues in the developing mouse conceptus. As noted previously, low but detectable amounts of *Ipl* mRNA are present in fetal liver and kidney, with *Ipl* mRNA more abundant than *Tih1* mRNA in fetal liver. Among the tissues examined in the adult mouse only the kidney shows *Ipl* expression at a level equal to or greater than *Tih1* (Fig. 4a, b).

We next asked whether the tissue-specific distributions of *Ipl* and *Tih1* mRNAs were paralleled by similar distributions of the corresponding proteins. This question is particularly relevant for an imprinted gene like *Ipl*, since imprinting of a protein-coding gene is expected to be effective as a mechanism of gene dosage regulation only if differences in mRNA levels translate into parallel differences in amounts of protein. To test this, we probed fetal and adult tissue lysates with polyclonal antibodies raised against IPL and TIH1 peptides. For both genes, the reactivity of the anti-peptide antisera against proteins of the predicted size verified the conceptual translation of the open reading frames. The Western blots showed that IPL and TIH1 migrate with approximate molecular masses predicted on the basis of sequence, 16.7 kDa and 13.7 kDa, respectively, and that the relative amounts of protein closely parallel the mRNA levels for both genes (Fig. 4b). Thus, both *Ipl* mRNA and protein are most abundant in mid-gestation extraembryonic membranes and placenta, are present in reduced amounts during later embryonic life, and are seen in adult tissues, at low levels, primarily in the kidney.

Expression of TIH1 protein continues throughout gestation, and in adult life is most prominent in those tissues, notably lung, which have high levels of message (Fig. 4b). Skeletal muscle, which is also strongly positive for *Tih1* mRNA, also expressed high levels of TIH1 protein, although the precise comparison of skeletal muscle to other tissues was not possible owing to difficulties in assessing the total protein loading for this actin-rich tissue (data not shown). These data suggest that, at least in non-stimulated tissues, neither *Ipl* nor *Tih1* is regulated significantly at the post-translational level. As a caveat, one family member, *TDAG51*, has been reported to be an inducible gene [activated by

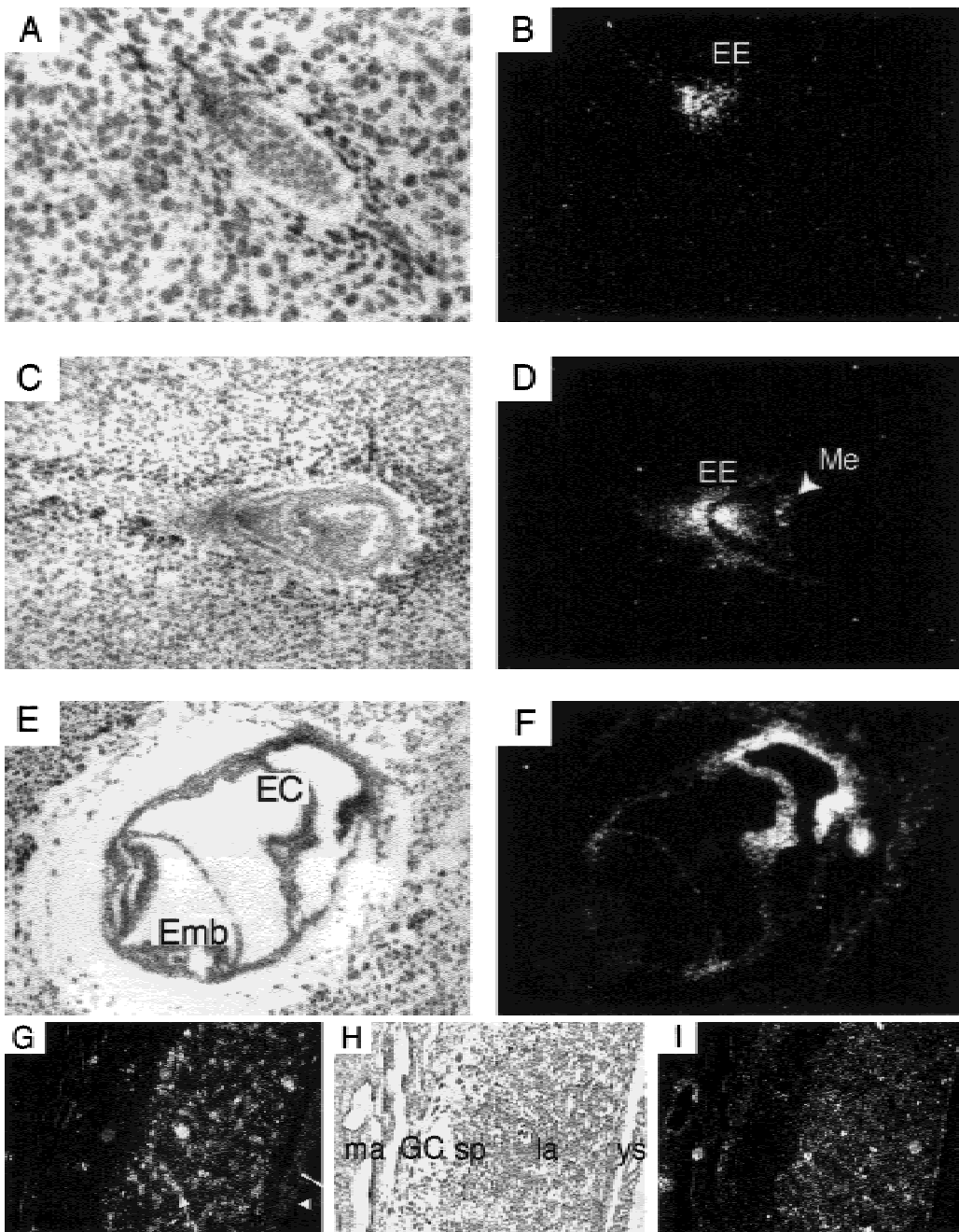
phorbol ester; (Park et al. 1996; Wang et al. 1998)], so it is possible that tissues which do not express particular family members in the non-stimulated state might do so under other environmental conditions.

We next used a human TIH1 cDNA probe to examine adult and fetal expression, and the findings paralleled those in the mouse tissue, with a broad range of tissues showing detectable expression, but with lowest expression in the liver and spleen (Fig. 4c). Interestingly, in an extension of our previous analysis of human IPL mRNA expression, we observed that a previously unrecognized site of high IPL expression is the adult prostate gland (Fig. 4c). In summary, the three members of this gene family show partially overlapping, but also distinct patterns of mRNA expression in mouse and human tissues, with *Ipl* mRNA being most restricted in its tissue distribution.

#### *Ipl* expression is cell type restricted in the extraembryonic tissues.

To gain insights at higher resolution, we examined the localization of *Ipl* mRNA and protein in early and mid gestation by RNA in situ hybridization (ISH) and immunohistochemistry. For the ISH studies we used both isotopic and non-isotopic methods, since these have complementary strengths (high sensitivity for isotopic and high resolution for non-isotopic). As detected by isotopic ISH, high levels of *Ipl* mRNA in the very early conceptus are limited to the polar trophectoderm. A strong signal is seen in the ectoplacental cone shortly after implantation at 5.5 days post-coitum (dpc), and high expression remains restricted to the extraembryonic tissues at later stages (Fig. 5). In contrast, sections from this early window of development (4.5–8.5 dpc) did not yield a detectable signal that *Tih1* mRNA (data not shown).

By 10.5 dpc and thereafter, both isotopic and non-isotopic ISH revealed high *Ipl* expression restricted to the labyrinthine trophoblast and the endodermal component of the visceral yolk sac (Figs. 5, 6). These two tissues are both extraembryonic, but are derived from different precursors: the polar trophectoderm and the primi-



**Fig. 5.** Isotopic ISH of *Ipl* and *Tih1* mRNAs in early murine development. Sections of developing mouse conceptuses were hybridized with the *Ipl* (B, D, F, G) or the *Tih1* (I) antisense cDNA probes. Bright-field (A, C, E, H) and dark-field (B, D, F, G, I) images are shown for stages 5.5 dpc (A, B), 6.5 dpc (C, D), 7.5 dpc (E, F), and 10.5 dpc (G, H, I). High expression of *Ipl* is seen in the implanting trophoblast at the earliest time point examined, and remains essentially restricted to the extraembryonic tissues at later stages. *Tih1* expression is not detectable at the early timepoints (data

not shown), but is seen diffusely in the labyrinthine trophoblast layer and the maternal tissues at 10.5 dpc (I). At this stage *Ipl* mRNA is seen in a more clustered pattern (arrow) and, in contrast to *Tih1*, is prominent in the yolk sac (arrowhead). EE = extraembryonic ectoderm, Me = primary mesoderm, EC = ectoplacental cavity, Emb = embryo, ma = maternal side of placenta, GC = giant cell layer, sp = spongiotrophoblast, la = labyrinthine trophoblast, ys = yolk sac.

**Fig. 6.** Non-isotopic ISH of *Ipl* mRNA in 11.5 dpc mouse tissues. Sections of whole conceptuses were hybridized with the digoxigenin-labeled *Ipl* cDNA antisense probe (A, C, D) or with the 4311 probe (Lescisin et al. 1988) (B), developed with alkaline-phosphatase anti-dig antibody and visualized either without counterstaining (A, B, D) or after a PAS counterstain (C). The specific *Ipl* signal is restricted to the labyrinthine trophoblast and visceral yolk sac endoderm, while the 4311 probe identifies the spongiotrophoblast layer. The control sense probe did not yield a signal above background (data not shown). Ma = maternal tissue, sp = spongiotrophoblast, la = labyrinthine trophoblast, pe = parietal endoderm, ve = visceral endoderm, emb = embryo, GC = giant cell layer of placenta.

**Fig. 7.** Immunohistochemical detection of IPL protein in 12.5 dpc mouse extraembryonic tissues. Sections of 12.5 dpc placenta and yolk sac were developed with anti-IPL antiserum (B, C, D) or with pre-immune serum (A). The immunoreactive *Ipl* protein is restricted to labyrinthine trophoblast and visceral yolk sac endoderm. Abbreviations as in Fig. 6, and fv = fetal vessel, bl = blood island, me = mesodermal surface of yolk sac.

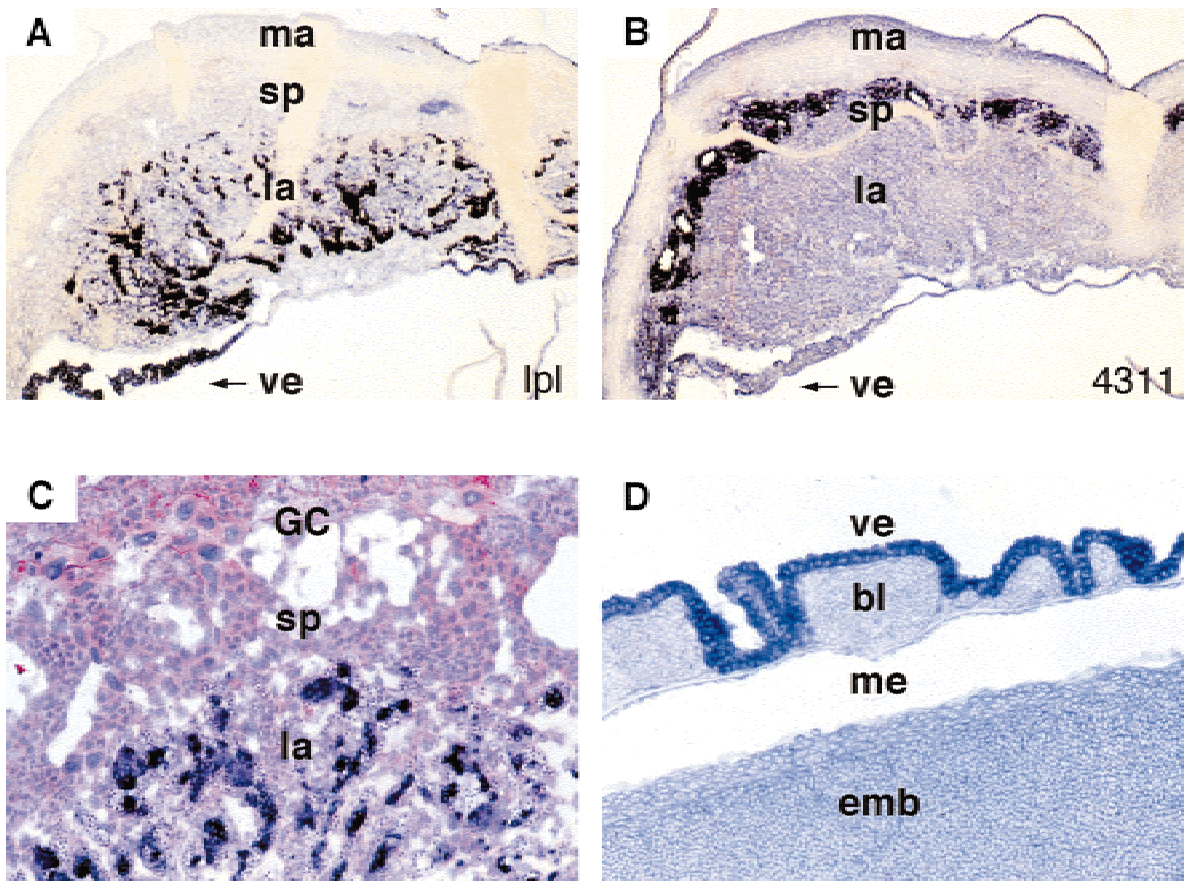


Fig. 6.

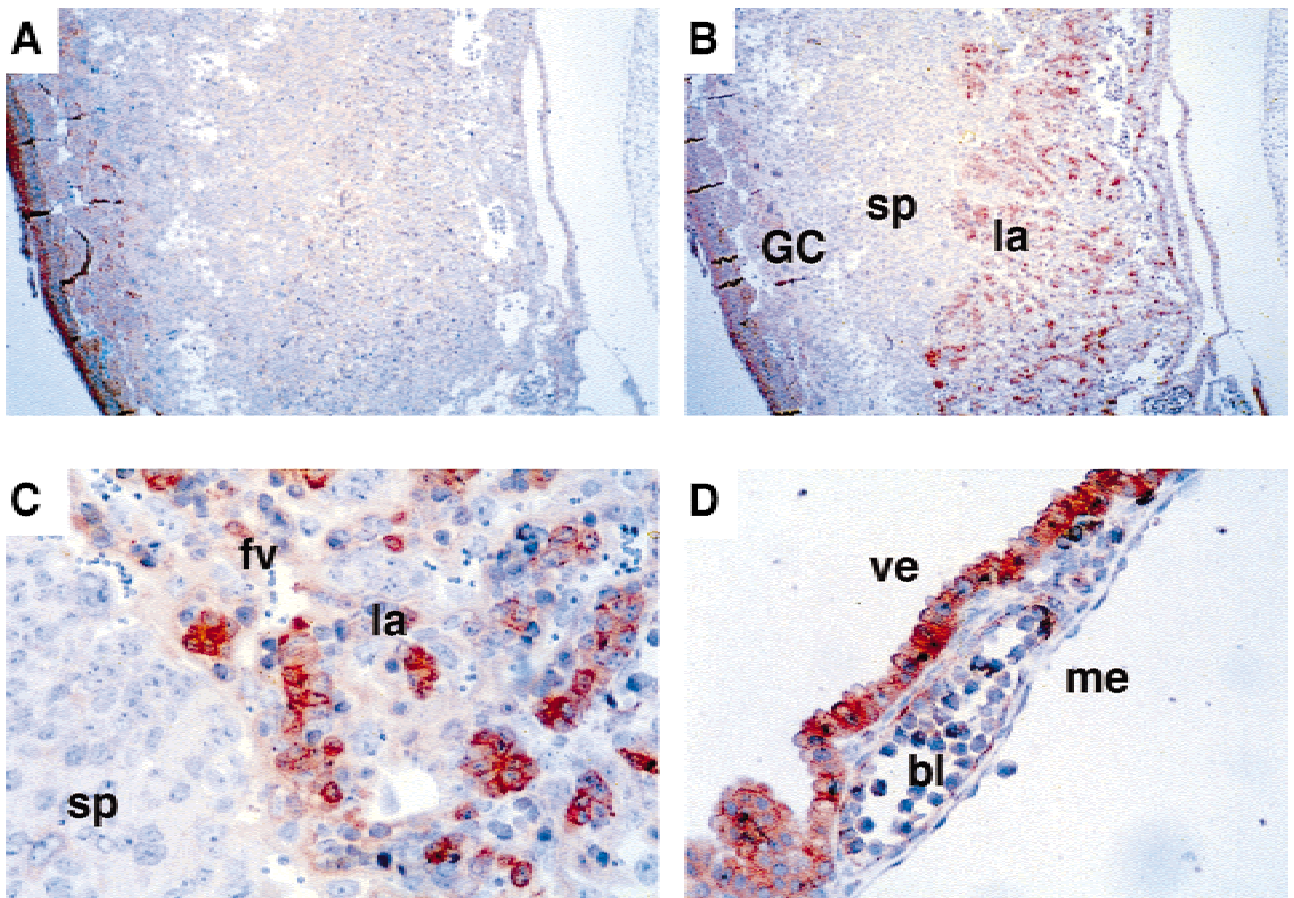


Fig. 7.



tive endoderm of the inner cell mass, respectively (Gardner 1983). No signal is observed in the histogenetically distinct spongiotrophoblast or giant cell layers of the placenta, or in mesenchymal cells of the placenta or yolk sac. The lack of *Ipl* expression in spongiotrophoblast was highlighted by hybridization of serial sections with a spongiotrophoblast-specific probe, the 4311 clone described by Rossant and coworkers (Lescisin et al. 1988). This yielded a "photographic negative" pattern of hybridization, compared with the *Ipl* pattern (Fig. 6).

Focality of the *Ipl* signal in the labyrinth was strongly suggested by the isotopic ISH and confirmed at higher resolution by non-isotopic ISH: this restriction of the signal to tight clusters of trophoblast cells became more pronounced with progression from the 10.5 to 14.5 dpc stages (Figs. 5, 6, and data not shown). The qualitative difference between embryonic (very low) and extraembryonic (very high) expression of *Ipl* is striking and precisely parallels the Northern blotting data in the current report and our previous study (Qian et al. 1997). The maternal components included in the sections were negative for detectable *Ipl* mRNA, confirming the validity of previous imprinting analysis of this gene in whole placenta.

By isotopic ISH the distribution of *Tih1* transcripts showed similarities and differences with *Ipl*. Consistent with the Northern analyses, *Ipl* mRNA appeared to be more abundant than *Tih1* mRNA in early (10.5 dpc) placenta, but the signal for *Tih1* mRNA became greater than for *Ipl* mRNA in later stage placenta (14.5 dpc; Fig. 5 and data not shown). *Tih1* mRNA was clearly detected in the labyrinthine trophoblast layer of the placenta, albeit in a more diffuse pattern than that of *Ipl*, but at no stage was the *Tih1* signal as strong as *Ipl* in the yolk sac (Fig. 5 and data not shown). Consistent with the Northern blotting results, in contrast to *Ipl*, a *Tih1* signal above background was clearly discernible in the maternal component of the placenta (Fig. 5 and data not shown).

The expression of *Ipl* specific to the labyrinthine trophoblast and visceral yolk sac was confirmed at the protein level by immunohistochemistry (Fig. 7). The pattern of IPL protein precisely paralleled that of its mRNA, with strong specific staining restricted to the labyrinthine trophoblast of the placenta and visceral endoderm cells of the yolk sac (compare Figs. 6, 7). Within the labyrinthine trophoblast, the staining was seen in tight clusters of relatively small cuboidal cells. Interestingly, in many areas these IPL-positive cells were separated by one cell layer from the adjacent blood vessel endothelial cells.

## Discussion

Genomic, gametic or parental imprinting is a unique mode of gene dosage regulation, and this epigenetic process influences various mammalian phenotypes, as well as susceptibility to neoplastic and non-neoplastic diseases in humans. Questions raised by current theories of imprinting include: (i) are there certain structural features that account for the targeting of some genes but not others for imprinting, and (ii) are there certain biological functions that are over-represented in the set of imprinted genes? In terms of the first sequence, there is statistical evidence suggesting that a compact structure (small genes with few introns) may be more common among imprinted genes than among non-imprinted ones (Hurst et al. 1996). However, in the well-characterized imprinted domain on mouse distal Chr 7/human Chr 11p15.5, there are examples of compact imprinted genes with few introns (*H19*, *Ipl*), moderately sized ones with more than 10 introns (*Impt1*), and very large ones with many large introns spanning more than 200 kb (*Kvlqt1*). Indeed, the very existence of imprinted domains argues strongly for chromosomal position as an important determinant of imprinting. The fact that *Ipl*, which closely resembles *Tih1* in sequence and exon/intron structure, is tightly imprinted in extraembryonic

and early fetal tissues, while *Tih1* is not imprinted in these or in any of multiple tissues examined at later developmental stages, suggests dominance of chromosomal context over exon/intron structure in genomic imprinting. Among gene pairs, very few others have been examined for concordant vs. discordant imprinting. The mouse *Ins1* and *Ins2* genes are small genes that share extensive sequence similarity. These genes are located on different chromosomes, and they are both functionally imprinted, at least in the yolk sac (Giddings et al. 1994). *Ins2* (mouse distal Chr 7) lies within a large imprinted chromosomal domain; whether *Ins1* (mouse Chr 19) is also within an imprinted domain is not known.

The second question deals with functional, rather than structural, aspects of imprinting and is motivated by the maternal-paternal conflict theory (Haig 1997; Haig and Graham 1991). This theory predicts that the set of imprinted mammalian genes should be enriched in examples that regulate fetal and placental growth. On the basis of its highly restricted regional expression, we suggest that the product of the imprinted *Ipl* gene may play a specific role in the two tissues that function in maternal/fetal nutrient and metabolite exchange, the visceral endoderm of the yolk sac and the labyrinthine trophoblast of the placenta. Another imprinted gene in the same chromosomal domain, *Mash2*, is specifically expressed in the placenta, and its germline deletion leads to a severe defect in placental development (Guillemot et al. 1995). We have made mice with a germline deletion of *Ipl*, and detailed characterization of the placentas in these animals is in progress, but it is already clear that these *Ipl*-mice are viable and fertile (D. Frank, T. Ludwig, and B. Tycko, unpublished data). The data in the current study, showing high sequence homology and overlapping tissue-specific expression patterns of the murine *Ipl* and *Tih1* genes, suggest that it may be necessary to simultaneously delete both of these genes to fully elucidate their function.

In addition to *Mash2*, at least four other genes in the *Ipl-H19* imprinted domain are highly expressed in the placenta, namely *Impt1/Orct12* (Dao et al. 1998; Cooper et al. 1998), *p57Kip2* (Lee et al. 1995), *Igf2* (Hedborg et al. 1994) and *H19* (Poirier et al. 1991; Walsh et al. 1995). Why are so many "placenta genes" clustered in a single ~1Mb imprinted chromosomal domain? Aside from chance, two possibilities occur a priori: there may be a structural explanation, that is, a shared long-range *cis*-acting element that promotes expression in the trophoblast, but alternatively there may have been a biological selection for imprinting of genes expressed in this extraembryonic organ.

Finally, the sequence data considered here define a novel gene family with a shared central amino acid sequence motif. We have used the term "PH domain-like structure" to describe this common motif in the *Ipl/Tih1/TDAG51* gene family; a definitive assignment of these shared sequences as standard PH domains will depend on structural studies. We suggest that in the future the nomenclature for the genes in this family might be usefully revised and unified to reflect the presence of this shared central motif.

**Acknowledgments.** This work was supported by grants from the National Institutes of Health to C. Mendelsohn, D. Frank and B. Tycko. We thank Giorgio Cattoretti for advice and assistance with immunohistochemistry.

## References

- Ainscough JF, Koide T, Tada M, Barton S, Surani MA (1997) Imprinting of *Igf2* and *H19* from a 130 kb YAC transgene. *Development* 124, 3621–3632
- Chaillet JR, Bader DS, Leder P (1995) Regulation of genomic imprinting by gametic and embryonic processes. *Genes Dev* 9, 1177–1187
- Chardin P, Paris S, Antony B, Robineau S, Beraud-Dufour S et al. (1996) A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains. *Nature* 384, 481–484
- Cooper PR, Smilnich NJ, Day CD, Nowak NJ, Reid LH et al. (1998)

- Divergently transcribed overlapping genes expressed in liver and kidney and located in the 11p15.5 imprinted domain. *Genomics* 49, 38–51
- Dao D, Frank D, Qian N, O'Keefe D, Vosatka RJ et al. (1998) IMPT1, an imprinted gene similar to polyspecific transporter and multi-drug resistance genes. *Hum Mol Genet* 7, 597–608
- Dao D, Walsh CP, Yuan L, Gorelov D, Feng L et al. (1999) Multipoint analysis of human chromosome 11p15/mouse distal chromosome 7: inclusion of H19/IGF2 in the minimal WT2 region, gene-specificity of H19 silencing in Wilms' tumor patients and methylation hyperdependence of H19 imprinting. *Hum Mol Genet* 8, 1337–1352
- Eck MJ, Dhe-Paganon S, Trub T, Nolte RT, Shoelson SE (1996) Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell* 85, 695–705
- Franco M, Boretto J, Robineau S, Monier S, Goud B et al. (1998) ARNO3, a Sec7-domain guanine nucleotide exchange factor for ADP ribosylation factor 1, is involved in the control of Golgi structure and function. *Proc Natl Acad Sci USA* 95, 9926–9931
- Gardner RL (1983) Origin and differentiation of extraembryonic tissues in the mouse. *Int Rev Exp Pathol* 24, 63–133
- Giddings SJ, King CD, Harman KW, Flood JF, Carnaghi LR (1994) Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting [see comments]. *Nat Genet* 6, 310–313
- Guillemot F, Caspary T, Tilghman SM, Copeland NG, Gilbert DJ et al. (1995) Genomic imprinting of Mash2, a mouse gene required for trophoblast development. *Nat Genet* 9, 235–242
- Haig D (1997) Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proc R Soc Lond B Biol Sci* 264, 1657–1662
- Haig D, Graham C (1991) Genomic imprinting and the strange case of the insulin-like growth factor II receptor. *Cell* 64, 1045–1046
- Hedborg F, Holmgren L, Sandstedt B, Ohlsson R (1994) The cell type-specific IGF2 expression during early human development correlates to the pattern of overgrowth and neoplasia in the Beckwith-Wiedemann syndrome. *Am J Pathol* 145, 802–817
- Hurst LD, McVean GT (1997) Growth effects of uniparental disomies and the conflict theory of genomic imprinting. *Trends Genet* 13, 436–443
- Hurst LD, McVean G, Moore T (1996) Imprinted genes have few and small introns [letter]. *Nat Genet* 12, 234–237
- Kolanus W, Nagel W, Schiller B, Zeitlmann L, Godar S et al. (1996) Alpha L beta 2 integrin/LFA-1 binding to ICAM-1 induced by cytohesin-1, a cytoplasmic regulatory molecule. *Cell* 86, 233–242
- Kozak M (1996) Interpreting cDNA sequences: some insights from studies on translation. *Mamm Genome* 7, 563–574
- Lee MH, Reynisdottir I, Massague J (1995) Cloning of p57KIP2, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution. *Genes Dev* 9, 639–649
- Lee MP, Feinberg AP (1998) Genomic imprinting of a human apoptosis gene homologue, TSSC3. *Cancer Res* 58, 1052–1056
- Lemmon MA, Ferguson KM (1998) Pleckstrin homology domains. *Curr Top Microbiol Immunol* 228, 39–74
- Lescisin KR, Varmuza S, Rossant J (1988) Isolation and characterization of a novel trophoblast-specific cDNA in the mouse. *Genes Dev* 2, 1639–1646
- Meacci E, Tsai SC, Adamik R, Moss J, Vaughan M (1997) Cytohesin-1, a cytosolic guanine nucleotide-exchange protein for ADP-ribosylation factor. *Proc Natl Acad Sci USA* 94, 1745–1748
- Mendelsohn C, Batourina E, Fung S, Gilbert T, Dodd J (1999) Stromal cells mediate retinoid-dependent functions essential for renal development. *Development* 126, 1139–1148
- Nagel W, Schilcher P, Zeitlmann L, Kolanus W (1998) The PH domain and the polybasic c domain of cytohesin-1 cooperate specifically in plasma membrane association and cellular function. *Mol Biol Cell* 9, 1981–1994
- Ohlsson R, Hedborg F, Holmgren L, Walsh C, Ekstrom TJ (1994) Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. *Development* 120, 361–368
- Park CG, Lee SY, Kandala G, Choi Y (1996) A novel gene product that couples TCR signaling to Fas(CD95) expression in activation-induced cell death. *Immunity* 4, 583–591
- Pfeifer K, Leighton PA, Tilghman SM (1996) The structural H19 gene is required for transgene imprinting. *Proc Natl Acad Sci USA* 93, 13876–13883
- Poirier F, Chan CT, Timmons PM, Robertson EJ, Evans MJ et al. (1991) The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. *Development* 113, 1105–1114
- Qian N, Frank D, O'Keefe D, Dao D, Zhao L et al. (1997) The IPL gene on chromosome 11p15.5 is imprinted in humans and mice and is similar to TDAG51, implicated in Fas expression and apoptosis. *Hum Mol Genet* 6, 2021–2029
- Rowe LB, Nadeau JH, Turner R, Frankel WN, Letts VA et al. (1994) Maps from two interspecific backcross DNA panels available as a community genetic mapping resource [published erratum appears in *Mamm Genome* 5:463, 1994]. *Mamm Genome* 5, 253–274
- Schimmoller F, Itin C, Pfeffer S (1997) Vesicle traffic: get your coat! *Curr Biol* 7, R235–R237
- Searle AG, Peters J, Lyon MF, Hall JG, Evans EP et al. (1989) Chromosome maps of man and mouse. IV. *Ann Hum Genet* 53, 89–140
- Shaw G (1996) The pleckstrin homology domain: an intriguing multifunctional protein module. *Bioessays* 18, 35–46
- Walsh C, Miller SJ, Flam F, Fisher RA, Ohlsson R (1995) Paternally derived H19 is differentially expressed in malignant and nonmalignant trophoblast. *Cancer Res* 55, 1111–1116
- Wang R, Zhang L, Yin D, Mufson RA, Shi Y (1998) Protein kinase C regulates Fas (CD95/APO-1) expression. *J Immunol* 161, 2201–2207
- Wu L, Ueda T, Messing J (1993) 3'-end processing of the maize 27 kDa zein mRNA. *Plant J* 4, 535–544