**Supplementary Information**

Table of Contents
1. SI Materials and methods
2. SI Figures
   a. SI Figure 1
   b. SI Figure 2
   c. SI Figure 3
   d. SI Figure 4
3. SI Tables
   a. Table I
   b. Table II
   c. Table III
   d. Table IV
   e. Table V
   f. Table VI
   g. Table VII
   h. Table VIII

**SI Materials and methods**

**Tissue Collection**

Mouse labyrinth tissue was micro-dissected from naturally mated crosses of C57/Bl/6J mice. Placentas were individually dissected on embryonic day 17.5 and the spongiontrophoblast layer was trimmed away using a number 10 scalpel blade. From each litter ¼ of the labyrinths were set aside for RNA extraction and microarray analysis and ¾ for sub-cellular fractionation and proteomic analysis. Total time of dissection until tissue homogenization in stabilizing buffer was less than 30 minutes. Human villous tree were dissected from term normal placenta delivered by cesarean section from a term pregnancy (~38 weeks). Total time from delivery to microdissection of villous trees and homogenization was 15-30 minutes.
Organelle Purification and Extraction

The preparation of mouse tissue organelle fractions was previously described (Kislinger et al., 2006). Briefly, mouse placenta labyrinth tissue was quickly excised, collected and washed several times in ice-cold phosphate-buffered saline to remove blood, and homogenized in ice-cold lysis buffer (250 mM sucrose, 50 mM HEPES (pH 7.4), 25 mM KCl, 5 mM MgCl2, 1 mM dithiothreitol (DTT), and 1 mM phenylmethylsulfonyl fluoride (PMSF) using a tight fitting glass pestle homogenizer. All subsequent steps were performed at 4°C or on ice. The lysate was centrifuged in a benchtop centrifuge at 800g for 10 min, with the supernatant serving as the source of cytosol, mitochondria and microsomes. The nuclear pellet was re-homogenized in lysis buffer, recentrifuged, solubilized in 5 ml of lysis buffer, and over-layered onto 4 ml of 0.9 M sucrose cushion buffer and centrifuged at 1,000g for 10 min in a swing-bucket rotor. The pellet was then re-solubilized in 5 ml of 1.9 M sucrose cushion buffer (1.9 M sucrose, 50 mM Tris-HCl (pH 7.4), 5 mM MgCl2, 1 mM DTT, and 1 mM PMSF). After over-laying onto 4 ml of 1.9 M cushion buffer, the nuclei were pelleted by ultracentrifugation at 150,000g for 1 h (Beckman SW 40Ti rotor). The mitochondria were isolated from the crude cytoplasm by centrifugation at 8,000g for 15 min, with the resulting pellet washed twice in lysis buffer. Mixed membranes were isolated by centrifugation of the post-mitochondrial cytoplasmic fraction at 150,000g for 1 h (Beckman SW 40Ti rotor), while the supernatant served as the cytosolic fraction (CYTO fraction). Nuclear proteins were extracted by re-suspending the nuclei in 5 volumes of nuclear extraction (NE) buffer (25 mM HEPES, pH 7.4, 400 mM NaCl), followed by 30 min incubation with gentle shaking at 4°C. The nuclei were then lysed by 10 passages through an 18-gauge needle. The debris was removed by
microcentrifugation at 20,000g for 30 min and the supernatant saved (NUC1 fraction). The pellet was resuspended in 5 volumes NET buffer (NE with addition of 1% Triton-X-100), gently shaken for 30 minutes 4°C, and the nuclei lysed by 10 passages through an 18-gauge needle. Debris was removed by microcentrifugation at 20,000g for 30 minutes, with the supernatant saved (NUC2 fraction). Mitochondrial proteins were isolated by incubating the isolated mitochondria in hypotonic lysis buffer (10 mM HEPES, pH 7.4, 1 mM DTT, 1 mM PMSF) for 30 min on ice. The suspension was briefly sonicated, shaken at 4°C for 1 h, and debris removed by microcentrifugation at 20,000g for 30 min. The supernatant was saved (MITO1 fraction). The pellet was resuspended in 5 volumes NET buffer (NE with addition of 1% Triton-X-100) and gently shaken for 30 minutes 4°C. Debris was removed by microcentrifugation at 20,000g for 30 minutes, with the supernatant saved (MITO2 fraction). Membrane proteins were extracted by re-suspending the microsomes in 5 volumes of NET buffer. The suspension was incubated with gentle shaking for 1 h. Insoluble membranes were removed by microcentrifugation at 20,000g for 30 min, while the supernatant was saved (MEMB fraction).

**In-solution Protein Digestion for MudPIT Analysis**

An aliquot of 150 μg of protein from each sample was precipitated overnight at -20°C with 5-volumes of ice-cold acetone, followed by centrifugation at 20,000 g for 15 min. The protein pellet was solubilized in 8M urea, 2 mM DTT, 50 mM Tris-HCl, pH 8.5 at 37°C for 1 hour, followed by carboxyamidomethylation with 10mM iodoacetamide for 1 hour at 37°C in the dark. The samples were then diluted with 100 mM Tris-HCl, pH 8.5 to ~1.5 M urea. Calcium chloride was added to a final concentration of 1mM and
digested with a 1:25 molar ratio of recombinant, proteomics grade trypsin (Roche Diagnostics, Laval, QC) at 37°C overnight. The resulting peptide mixtures were solid phase-extracted with Varian OMIX cartridges (Mississauga, ON, Canada) according to the manufacturers instructions and stored at -80°C until further use.

**MudPIT Analysis**

A fully automated 9-cycle, 18-h MudPIT procedure was set up similar as previously described (Kislinger et al., 2006; Washburn et al., 2001). A quaternary HPLC-pump was interfaced with a LTQ linear ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with a nano-electrospray source (Proxeon Biosystems, Odense, Denmark). A 100 μm inner diameter fused silica capillary (InnovaQuartz, Phoenix, AZ) was pulled to a fine tip using a P-2000 laser puller (Sutter Instruments, Novato, CA) and packed with ~7 cm of Jupiter™ 4μ Proteo 90Å C12 reverse phase resin (Phenomenex, Torrance, CA), followed by ~5 cm of Luna® 5μ SCX 100Å strong cation exchange resin (Phenomenex, Torrance, CA). Samples were loaded manually on separate columns using an in-house pressure vessel. As peptides eluted from the microcapillary columns, they were electrosprayed directly into the MS. A distal 2.3 kV spray voltage was applied to the microsplitter tee (Proxeon Biosystems). The MS operated in a cycle of one full-scan mass spectrum (400-1400 m/z), followed by 6 data-dependent MS/MS spectra at 35% normalized collision energy, which was continuously repeated throughout the entire MudPIT separation. The MS functions and the HPLC solvent gradients were controlled by the Xcalibur data system (Thermo Fisher Scientific, San Jose, CA).
**Protein Identification**

Raw files were converted to m/zXML using ReAdW and searched by X!Tandem against a mouse (v3.28) or human (v3.28) IPI (International Protein Index; [http://www.ebi.ac.uk/IPI](http://www.ebi.ac.uk/IPI)) protein sequence database. To estimate and minimize our false positive rate the protein sequence database also contained every IPI protein sequence in its reversed amino acid orientation (*target-decoy strategy*) (Gortzak-Uzan et al., 2008; Sodek et al., 2008). A conservative false discovery rate was set to 0.5% on the peptide level, as recently described (Gortzak-Uzan et al., 2008; Sodek et al., 2008). Only fully tryptic peptides matching these criteria were accepted to generate the final list of identified proteins. We only accepted proteins identified either with two unique peptides or one unique, unambiguous peptide found in ≥2 MudPIT runs per analyzed organelle fraction. To minimize protein inference, we developed a database grouping scheme and only report proteins with substantial peptide information, as recently reported (Gortzak-Uzan et al., 2008; Sodek et al., 2008).

**Western Blot Analysis**

Approximately 50 ug of protein from each sample was denatured in loading buffer, boiled for 10 minutes and loaded onto 7.5% or 10% SDS-PAGE gels. Proteins were separated by electrophoresis and transferred to Nitrocellulose membrane (Bio-Rad Laboratories, Mississauga, ON). Epitopes were detected with antibodies against eNOS (Transduction Labs), F1-ATP Synthase-α (gift of Dr. Gramolini Lab, University of Toronto), Transketolase (gift of Dr. Schimmer, University of Toronto) or Histone H3 (Abcam). After several washes in TBST, membrane blots were incubated with either goat
anti-rabbit IgG horseradish peroxidase (Sigma) or goat anti-mouse IgG peroxidase (Sigma) secondary antibody, diluted at 1:10,000 in TBST, for 1 h at room temperature. Following several washes in TBST, the membranes were developed with SuperSignal West Pico Chemiluminescent Substrate (Pierce, Rockford, IL) according to the manufacturer’s recommendation.

**Comparison to Other Proteomics Datasets**

Comparison of our labyrinth proteome resource to other recently published proteome profiles (Kislinger et al., 2006; Adachi et al., 2007; Foster et al., 2006; Gramolini et al., 2008; Graumann et al., 2008) was achieved using the ProteinCenter bioinformatics software (Proxeon Biosystems, Odense, Denmark). To link these data sets, we loaded the accession keys for each of these projects into ProteinCenter and BLASTed them against each other. Only proteins with at least 95% sequence homology were considered identical.

**Microarray Analysis**

Total RNA was extracted from the pooled labyrinths using the Trizol reagent. A total of four litters were used. Balanced aliquots from each litter were then pooled into two technical replicate samples for microarray analysis on Affymetrix MOE430v2 microarrays. Data was processed in GCOS (version 1.4.0.036) for expression analysis for absent/present calls and signal intensity. Four villous trees from each of two individual placentas were separately extracted for total RNA using Trizol reagent (Invitrogen). Two
technical replicate samples were prepared by using a balanced aliquot from each of the eight samples. The replicates were analyzed on Affymetrix HG133plus microarray. Data was processed in GCOS (version 1.4.0.036) for expression analysis for absent/present calls and signal intensity.

**Protein to Microarray Comparison**

Annotation tables were obtained from Ensembl using BioMart (www.ensembl.org/biomart/martview) with annotation of gene IDs, IPI IDs and Affymetrix probe IDs. MySQL database tables of protein data with IPI IDs, microarray data with Affymetix probe IDs and Biomart annotation tables were generated. SQL commands were used to link the protein and microarray datasets and add Ensembl gene IDs.

**Ortholog Comparisons**

A table of human-mouse gene orthologs was obtained from Ensembl using BioMart (www.ensembl.org/biomart/martview) along with annotation for the ortholog mapping (one-to-one, one-to-many and many-to-many), and ortholog scores. These were then added to the MySQL database of protein and microarray data. SQL commands were used to link the orthologs and generate a gene ID centered non redundant list of mouse to human orthologs and their corresponding protein and microarray data. The redundant data set was sorted such that the IPI ID with the largest number of detected spectral counts was used as the representative IPI for the Ensembl gene and all other matching IPI
IDs were then removed. The microarray probe sets were sorted for a present call. If any probe set had a present call linked to a gene ID it was used as the representative probe set ID and all others were removed. If no probe set had a present call then a randomly chosen probe set was used. In cases where no protein data existed null values were inserted into the table.

**Jaccard Score**

A Jaccard score was generated by the comparison of the detection call generated by GCOS analysis. P, M and A values were transformed into 1 for all P and 0 for M and A. the Jaccard score is the comparison of all attributes (genes) for which there was a measurement. Any attribute for which there was no measurement in either condition is ignored. The Jaccard score, $S_{ij}$ for any two sets (i and j) is given by the equation

\[
S_{ij} = \frac{a_{ij}}{a_{ij} + b_{ij} + c_{ij}}
\]

where a, b and c are the sum of all attribute values of (1, 1), (1,0) and (0, 1). In this way a score of 1 is a perfect correlation and a score of 0 is no correlation.

**Gene Ontology**

Calculation of enrichment of Gene Ontology terms was done using the ArrayTrack library (www.fda.gov/nctr/science/centers/toxicoinformatics/ArrayTrack/index.htm). IPI IDs from selected clusters were entered into the GOFFA library, which then calculated a Fisher exact P-value and fold enrichment versus the entire database.
Protein-Protein Interaction Network Analysis

Mouse protein-protein interactions were obtained from the Interologous Interaction Database Ver. 1.71 (I2D; http://ophid.utoronto.ca/i2d; Aug. 22, 2008). I2D contains interactions for human, rat, mouse, fly, worm and yeast from three sources: low throughput interactions compiled from major protein-protein interaction databases that curate literature, high-throughput interactions from individual studies and interologous predicted interactions, generated by mapping known interactions from all species to orthologous protein pairs in other species. All three sources of interactions were used; the only restriction was that interologous interactions were based only on interactions from human and rat. The SwissProt IDs of interacting proteins were mapped to Ensembl gene IDs using BioMart (www.ensembl.org/biomart/martview; Aug. 22, 2008). When several SwissProt IDs mapped to a single Ensembl gene, all the interactions of these SwissProt IDs were assigned to a single ‘gene product’ represented by the Ensembl gene ID. In this sense we did not distinguish specific splice variants. The resulting network contained 49,509 interactions among 9,355 Ensembl genes. From this network, a sub-network was selected comprised of genes with one-to-one human orthologs, that had mRNA and protein detected in both mouse and human. This sub-network, $G$, contained 6,233 interactions among 1,762 genes.

In the sub-network we identified a set of genes, $S$, with known placental labyrinth/vasculature phenotypes, and looked for genes enriched for interactions with this set. First, we identified 23 genes with the phenotypes *abnormal placental labyrinth morphology* (MP:0001716) or *abnormal placental vasculature* (MP:0003231), based on annotations in the Mouse Genome Informatics (MGI; http://www.informatics.jax.org/)
database (Aug. 22, 2008). Next, we identified genes significantly enriched for interactions with this set, S. We considered enrichment from the following perspective: if a gene had $n$ interactions with set $S$ how often would it have as many interactions with a randomly chosen set of genes from our sub-network? To implement this approach we chose 10,000 random sets of genes from our sub-network. Each set, $S_i$, had the same number of genes as $S$, and had a similar degree distribution (i.e., if the top $i\%$ of degrees in $S$ were in the range $[a,b]$, the top $i\%$, of degrees in $S_i$ were in the same range; with $i$ restricted to single percent increments). For each sub-network gene not in $S_i$ or $S$, the number of interactions with $S_i$ was determined. Since the number of interactions with $S_i$ depends largely on degree, we divided genes into groups based on their degree (i.e., if a degree, $k$, occurred for less than 1% of the genes in the sub-network, $k$ was included in a range of degrees that occurred in at least 1% of genes). We then compiled a table of interaction frequencies, indicating how frequently a gene of degree, $m$, had $n$ interactions with randomly chosen sets. Based on these frequencies we calculated the probability of a gene with degree, $m$, having at least $n$ interactions with $S$. There were 34 genes with probabilities $\leq 0.05$. These genes had interactions with 15 of the genes from set $S$.

References


Validation of fraction purity. Shown are heat maps for all proteins identified in human and mouse demonstrating the separation of isolated organelle fractions. Western blotting and Gene Ontology analyses independently confirmed the enrichment of our isolated fractions. Note that the proteomic data for NUC-1 and NUC-2 has been merged in the heat maps but is shown immunoblotted separately.
Percentage of all proteins identified and the total number of replicates in which they were observed. Approximately 50% of the proteins from any single cell fraction (see legend) were observed in all 5 replicates and greater than 95% of all proteins were observed in 2 or more replicates. This indicates that the data set is near saturation of detection for the analytical methods employed in this study. cyto, cytosol cell fraction; memb, Triton X-100 soluble microsome cell fraction; mitoi, soluble mitochondrial fraction; mitoii, Triton X-100 soluble mitochondrial fraction; nuci, high salt soluble nuclear fraction; nucii, Triton X-100 soluble nuclear fraction.
**In silico analysis of peptides.** (A) Peptides in cluster I and III were in silico digested to generate tryptic peptides. (B) Peptide masses of generated peptides. (C) Hydrophobicity of generated peptides. (D) Isoelectric point of generated peptides.
Immunohistochemistry images of human placenta from Protein Atlas that were scored as no expression. Itgb3 and DCN have clear expression in the villous mesenchyme while Icam1 is only expressed in the endothelium. Itga2b and Cd82 appear to have no expression in the villous structure (not shown for Cd82), however Cd82 did have expression in the decidua (shown). Images modified from Protein Atlas (Ponten et al., 2008; www.proteinatlas.org)