UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:
Questa è la versione dell'autore dell'opera:
[Hum. Mol. Genet. (2015) 24 (1): 118-127.
doi: 10.1093/hmg/ddu423]

The definitive version is available at:
La versione definitiva è disponibile alla URL:
[http://hmg.oxfordjournals.org/content/24/1/118.long]

Susceptibility allele-specific loss of miR-1324-mediated silencing of the *INO80B* chromatin-assembly complex gene in pre-eclampsia

Cees B.M. Oudejans[1,4,*]**, Omar J. Michel**[1]**, Rob Janssen**[2,4]**, Rob Habets**[1]**, Ankie Poutsma**[1]**, Erik A. Sistermans**[3]**, Marjan M. Weiss**[3]**, Danny Incarnato**[5,6]**, Salvatore Oliviero**[5,6]**, Gunilla Kleiverda**[7]**, Marie Van Dijk**[1,4] **and Reynir Arngrímsson**[8]

[1]Department of Clinical Chemistry
[2]Department of Physiology
[3]Department of Clinical Genetics and
[4]Institute for Cardiovascular Research, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands
[5]Human Genetics Foundation (HuGeF), Via Nizza 52, 10126 Torino, Italy
[6]Dipartimento di Biotecnologie, Chimica e Farmacia, Università degli Studi di Siena, Via Fiorentina 1, 53100 Siena, Italy
[7]Department of Gynecology, Flevo Hospital, Hospitaalweg 1, 1315 RA Almere, The Netherlands and
[8]Biomedical Center, University of Iceland, Vatnsmýrarvegi 16, 101 Reykjavík, Iceland

Abstract

In humans, the elucidation of the genetics underlying multifactorial diseases such as pre-eclampsia remains complex. Given the current day availability of genome-wide linkage- and expression data pools, we applied pathway-guided genome-wide meta-analysis guided by the premise that the functional network underlying these multifactorial syndromes is under selective genetic pressure. This approach drastically reduced the genomic region of interest, i.e. 2p13 linked with pre-eclampsia in Icelandic families, from 8 679 641 bp (region with linkage) to 45 264 bp (coding exons of prioritized genes) (0.83%). Mutation screening of the candidate genes (n = 13) rapidly reduced the minimal critical region and showed the INO80B gene, encoding a novel winged helix domain (pfam14465) and part of the chromatin-remodeling complex, to be linked to pre-eclampsia. The functional defect in placental cells involved a susceptibility allele-dependent loss-of-gene silencing due to increased INO80B RNA stability as a consequence of differential binding of miR-1324 to the susceptibility allele of rs34174194. This risk allele is located at position 1 in an absolutely conserved 7-mer (UUGUCUG) in the 3-UTR of INO80B immediately downstream of a variant Pumillio Recognition Element (UGUANAAG). These data support that pre-eclampsia genes affect a conserved fundamental mechanism that evolved as a consequence of hemochorial placentation. Functionally, this involves founder-dependent, placentally expressed paralogous genes that regulate an essential trophoblast differentiation pathway but act at different entry points.

INTRODUCTION
The elucidation of the genetics of common genetic traits remains complex. For pre-eclampsia, de novo hypertension with proteinuria during pregnancy, the existence of multiple susceptibility loci (4q34, 2p13, 9p13, 2p25, 10q22, 12q23) in different founder populations (Australia, Iceland, Norway, Finland, Netherlands) is known for decades, but with limited success in the identification of the genes involved (1,2). Two pre-eclampsia genes have been identified in Dutch pre-eclamptic females (1,2). At least several additional loci are predicted to exist in other populations (3) (Table 1). Here, we demonstrate that pathway-guided genome-wide meta-analysis permits rapid identification of susceptibility genes as tested in the Icelandic population.

Table 1.
Loci with genome-wide significant linkage to pre-eclampsia in populations other than the Dutch

This approach is based on the premise that the functional network underlying these traits, despite their segregation with different populations, is the same and under selective pressure. We recently showed that the Hemolysis, Elevated Liver enzymes, Low Platelets (HELLP) syndrome, a severe variant of pre-eclampsia, is caused by a large intergenic non-coding RNA, LINC-HELLP, expressed in the placenta and activating at least 1198 effector genes mainly involved in the cell cycle (2). Dysfunction of this gene leads to defective trophoblast invasion (2). In the extravillous trophoblast, the cell central in the etiology of pre-eclampsia, cell cycle regulation is intimately linked to the process of invasion into the maternal tissue and concomitant maternal vessel wall modification. Proliferative, non-invasive villus trophoblast cells from the anchoring villi differentiate into non-proliferative, invasive trophoblast cells, migrate away from the villus (hence called extravillous) toward and into the maternal spiral arteries and transform these into high conductance, low resistance arteries. During this process, the endovascular extravillous trophoblasts cells become embedded in the transformed arteries, while the so-called interstitial trophoblast cells located in the maternal placental bed differentiate into polyploid giant cells. In the mouse, the giant cells are invasive; in humans, the giant cells are non-invasive, end-stage differentiated cells. This interaction between fetal and maternal cells leads to the establishment of an essential connection between the maternal and fetal circulations that meets the demands of the growing fetus for its supply of oxygen and nutrients throughout pregnancy. Failure of this process by trophoblast dysfunction, as central in the etiology of pre-eclampsia, leads to incomplete vessel wall modification triggering a sequel of events that ultimately, months after the initial etiological event, leads to maternal symptoms of hypertension and proteinuria. In other words, the placental genotype controls the maternal phenotype. This paradox is central in the familial, genetic forms of pre-eclampsia that present with early-onset symptoms in the mother (< week 34) and intra-growth restriction of the fetus and carry the highest risk of maternal and fetal morbidity and mortality.

We argued that this (dys)functional network in the extravillous trophoblast involves additional susceptibility genes for pre-eclampsia in different founder populations if these genes are: (i) located in loci with confirmed genome-wide linkage (e.g. 2p13), (ii) transcribed in the cell central in the etiology (i.e. extravillous trophoblast) and (iii) involved in the same pathway (i.e. trophoblast invasion linked with cell cycle exit) (1,2,4). We used these criteria for qualitative analysis of the genome-wide linkage- and expression scans available for pre-eclampsia to filter the pooled data allowing selective targeting of candidate genes for pre-eclampsia.

RESULTS
Pathway-guided genome-wide meta-analysis leads to a drastic reduction of the genomic region of interest for candidate gene studies
We tested this approach for the 2p13 locus linked with pre-eclampsia in Icelandic families (criterium 1). We used the transcriptome profile of the SGHPL5 extravillous trophoblast cell line, the cell central in the etiology of pre-eclampsia (criterium 2). The expression profile was obtained at high resolution following controlled experimental disruption of cell cycle exit as essential for the end-stage differentiation of these cells (criterium 3) (2,4). The latter effect was seen after knockdown of the LINC-HELLP RNA (2). Transcripts, both primary and isoforms, differentially expressed in extravillous trophoblast cells (SGHPL5) after siRNA-mediated knockdown of the HELLP large intergenic noncoding RNA (lincRNA) (n = 5041, q-value ≤ 0.05) were screened for their location in chromosomal loci with confirmed linkage to pre-eclampsia. In the 2p13 locus (8 679 641 bp between D2S292-D2S329) linked with pre-eclampsia in Iceland (lod 4.77) and Australia (lod 3.5) (4,5), 122 genes are expressed in the extravillous trophoblast. Compared to the expression in all tissues, the trophoblast genes are restricted to the telomeric half of the locus with the 3′-end boundary marked by C2orf3/MRPL19. Of these, 11 annotated genes, ANKRD53, MCEE/MPHOSP10, ZNF638, DYSF, ALMS1/ALMS1P, TPRKB, BOLA3/MOBKL1B, TET3, INO80B/WBP1, HK2, MRPL19/GCFC2 and 2 novel genes **(**Fig. 1) were selected for mutation analysis as they met the third criterium as well (Supplementary Material, File S1). By this pathway-guided, genome-wide meta-analysis, the region of interest for screening of pre-eclampsia candidate genes was reduced drastically from 8 679 641 bp (region with linkage) to 45 264 bp (coding exons of prioritized genes) (0.83%).

Pathway-guided genome-wide meta-analysis targets the INO80B gene as pre-eclampsia gene. The 2p12-13 locus (8 679 641 bp between D2S292-D2S329) linked with familial pre-eclampsia in Iceland (lod 4.77) and Australia (lod 3.5) contains 122 genes (**A**) expressed in the extravillous trophoblast. Compared with the expression in all tissues (**C**), the trophoblast genes are restricted to the telomeric half (A). Of these, 17 genes are involved in the same functional network underlying pre-eclampsia or HELLP (Supplementary Material, File S1). Thirteen genes (**B**) were selected for sequencing of informative affected females (Supplementary Material, File S2). This approach reduced the genomic region of interest to 0.83% and identified a minimal critical region consisting of a solid spine of LD block (triangle) with the minor alleles of the first (rs2272240) and last (rs2306806) markers identical in all patients (Fig. 2) and containing the INO80B-WBP1 gene.

The INO80B-WBP1 gene is a susceptibility gene for pre-eclampsia
The coding regions of these prioritized genes were subsequently genotyped in a set of five informative affected women (Supplementary Material, File S2). By this, 75 SNPs were found. The identity-by-descent allele-sharing pattern of these identified two solid spine of linkage disequilibrium (LD) blocks with the minor alleles of the outer markers shared in all affected women. By definition, the intermediate markers are not necessarily in LD with each other in these blocks. An upper block (chr2:71 209 282–71 838 597) (629 316 bp) between ANKRD53-DYSF consisting of two smaller LD blocks and a second, lower block (chr2:74 377 419–75 113 572) (736 154 bp) bordered by BOLA3/AS1-HK2 containing three smaller LD blocks were seen (Fig. 2). To select between these, we searched for functional and other features of the genes within these blocks to see if they were similar or identical to other genes linked with familial early-onset pre-eclampsia. The highest and in fact the only similarity was seen for the INO80B-WBP1 gene. Besides being located in the minimal critical region, it resides within the EGR4-TET3-INO80B-DQX1-HK2-TACR1 region paralogous to the EGR2-TET1-STOX1-DDX50-HK1-TACR2 region on 10q22. The latter contains the winged helix domain STOX1 gene causing pre-eclampsia in Dutch females and transgenic mice (1,6). This criterium excluded the upper ANKRD53-DYSF block and its associated genes. Secondly, the INO80 chromatin-assembly complex, known to be involved in cell cycle regulation, also controls polyploidy (7). Both features are central in pre-eclampsia and HELLP (1,2). In accordance, both INO80B and STOX1A proteins bind to the microtubule (8,9). Finally, one of the eight INO80 family members, INO80G (NFRKB) was recently shown to encode a novel winged helix domain (pfam14465) (10). By secondary structure prediction (PsiPred), we found the INO80B protein (NP_112578) to contain this novel winged helix domain as well (Supplementary Material, File S3).
The INO80B-WBP1 gene is a susceptibility gene for pre-eclampsia. The identity-by-descent allele sharing pattern of 75 SNPs identified 2 solid spine of LD blocks (indicated by]). The upper block (1) is located between ANKRD53-DYSF. The lower block (2) is bordered by BOLA3/AS1-HK2. Using the additional criteria of paralogous region, cell cycle control and winged helix domain (see main text for

explanation), only the 707 620 bp region (3) containing the INO80B-WBP1 gene qualified as minimal critical region. In this region, the minor alleles of the first (rs2272240) and last (rs2306806) markers identical in all five patients. The minor alleles are indicated in gray. The risk allele is indicated in red. The regions identical-by-descent are indicated in yellow.

In-depth mutation analysis, both qualitative and quantitative, was therefore performed for the INO80-WBP1 gene. Using high-density genome-wide scans, we excluded quantitative mutations such as insertions or deletions in INO80B as cause of disease in these pre-eclamptic females (Supplementary Material, File S4). The array used, CytoScan HD, has greater than 99% sensitivity and can reliably detect 25–50 kb copy number changes across the genome at high specificity with SNP (allelic) call corroboration. No disease-causing quantitative aberrations were found. Only regions with loss-of-heterozygosity (LOH) were found being present in unrelated individuals as well. Interestingly, this LOH region coincided with the region subject to strong positive natural selection in individuals of European descent as identified by Grossman using the Composite of Multiple Signals (CMS) approach (11) (Supplementary Material, File S5). Although excluding quantitative mutations as cause of pre-eclampsia in the families we tested, the LOH and CMS findings once again support our previous notion that the regions harboring pre-eclampsia genes appear subject to conserved evolutionary pressure. This can extend to other organs (brain) and diseases (Alzheimer disease) and present as shared pathways such as the amyloid pathway as we previously showed (9) and as was recently confirmed by Buhimschi and co-workers (12).

Following sequence analysis of the transcript (Supplementary Material, Files S6A and B) and transcription unit (chr2:74681496–74688014) of INO80B-WBP1, rs34174194 was the only variation that qualified as a susceptibility allele. The minor allele (G) is present in all patients, while the major allele (T) is absolutely conserved at the DNA level between species as remote as Opossum (Supplementary Material, File S7).

Using the high-resolution RNA sequence data (2), the INO80B-WBP1 readthrough transcript containing the novel winged helix domain encoded by INO80B was the only expressed in extravillous trophoblast. This transcript corresponds to cuff.25295.1 (Supplementary Material, File S6A). Its expression level [Fragments Per Kilobase Mapped (FPKM) value of 32.9] correlates with medium (FPKM 7.5–29.99) to high (FPKM > 30) abundancy expression levels (13). For comparison, STOX1 expression levels in these cells are low (FPKM 0.93). This indicated that rs34174194 forms part of the INO80B transcript rather than forming part of the upstream region of the WBP1 gene. We confirmed this using strand-specific RT–PCR: rs34174194 is located in the 3′-untranslated region (3UTR) of the INO80B-WBP1 transcript (data not shown). This transcript is 2638 bases in size and encodes a protein (NP_112578) of 356 amino acids with a predicted molecular weight (MW) of 38638 and isoelectric point (pI) of 9.66 (Supplementary Material, File S6B).

Susceptibility allele-specific loss of miRNA-mediated silencing of INO80B chromatin-assembly complex transcripts

We noted striking parallels with a previous study (14) that could explain the genotype-phenotype correlation. In the development of the vertebrate nervous system, a distinctive step occurs at mitotic exit when cells lose multipotency and begin to develop stable connections. A switch in ATP-dependent chromatin-remodeling accompanies this transition. Functionally, this switch is microRNA-mediated involving ultraconserved sequences in the 3UTR of chromatin-assembly complex proteins. These ultraconserved sequences correspond to recognition sites for specific microRNAs with restricted expression (14).

We analyzed the INO80B transcript for the presence of ultraconserved sequences and microRNA binding sites using MREdictor, PITA and TargetScan (15–17). Two highly conserved 7-mers (UUGUCUG) flanking a variant Pumillio Recognition Element (PRE) (UGUAnAAG) were found within a conserved 70 bp region of the INO80B 3UTR (Fig. 3). In the presence of the susceptibility allele (G at position 1 of the second UUGUCUG site), seed pairing with miR-1324 would change binding to the canonical 7mer-m8 site (seed match + position 8) with strong repression efficiency into binding to a marginal 6mer site (seed match limited to positions 2–7) with marginal repression (18,19). Hence gene silencing of INO80B would be less sufficient or incomplete for the genotype associated with pre-eclampsia. The secondary structures (RNAfold) of the conserved sequences in the INO80B-3UTR showed the formation of hairpins with the sequence containing rs34174194 in a single-stranded, i.e. accessible conformation for both wild-type and susceptibility alleles (Supplementary Material, File S8). Two additional microRNAs were predicted to bind differentially to the region including the susceptibility allele: miR-409-5p and miR-574-3p. Similar to miR-1324, miR-409-5p binding would lead to increased INO80B levels in the presence of the G-allele, while miR-574-3p would lead to decreased INO80B levels in the presence of the G-allele. However, binding by miR-1324 was the only

combination where the position of rs34174194 directly involves and affects pairing with the seed sequence (CAGACAG for miR-1324). Binding by miR-409-5p and miR-574-3p would not involve a susceptibility allele-dependent direct effect on seed pairing but affect secondary structure.

The 3UTR region of INO80B containing the risk allele for pre-eclampsia is highly conserved and predicted to bind differentially to miR-1324. The conserved 70 bp region of the 3UTR of INO80B contains two highly conserved 7-mers (UUGUCUG) (sites 1 and 2) flanking a variant PRE (UGUAnAAG). Both microRNA target sites are at cooperative distance (>7 and < 40 nucleotides). MicroRNA-target interaction prediction as done with PITA, MREdictor and TargetScan indicated differential binding of miR-1324, miR-409-5p and miR-574-3p to the second site. miR-1324 and miR-409-5p would bind preferentially to the transcript carrying the wild-type allele (T) and miR-574-3p to the transcript carrying the risk-allele (G). Only for miR-1324, this would involve a direct effect on seed pairing changing the canonical 7mer-m8 target site (with strong repression) to a marginal 6 mer target site (with weak repression). For the other two microRNAs (miR-409-5p, miR-574-3p), this would involve an (indirect) effect on secondary structure by a change in duplex minimal free energy (MFE) (ΔG_{duplex}) and accessibility cost (ΔG_{access}).

For the microRNA-chromatin switch to exist in vivo, trophoblast cells should express these microRNAs. MiR-574-3p and miR-409-5p are preferentially expressed in the placenta, the latter being part of the imprinted 14q32 domain (20,21). The expression of miR-1324 appears highly restricted. By microRNA cloning, miR-1324 was identified in neuroblastoma stage 4S, a rare entity, also called 'special' neuroblastoma (22). By deep sequencing, miR-1324 was also identified in squamous cells (about 83 reads per million) (23). By stemloop RT-PCR, we confirmed that miR-1324 is expressed in extravillous trophoblast cells (Supplementary Material, File S9).

Using luciferase reporter assays, we tested whether microRNA-mediated regulation of the INO80B chromatin assembly complex protein is a feature in trophoblast cells and whether gene silencing is affected when the susceptibility allele is expressed. Under default conditions (no microRNA inhibitor present) (Fig. 4A), the levels of the INO80B transcript carrying the susceptibility allele (G) remained significantly higher in SGHPL5 cells than the levels of the INO80B transcript carrying the wild-type allele (T). This excluded miR-574-3p as a mediator of the predicted differential effect as in that case the levels of T and G alleles would be opposite to the levels observed. After microRNA inhibition with anti-miR-409-5p, the levels of the INO80B transcript carrying the susceptibility allele (G) remained similar to the default situation with significantly higher levels of the INO80B transcript carrying the susceptibility allele (G) under all concentrations tested (Fig. 4B). In contrast, during inhibition with anti-miR-1324, the INO80B transcript levels carrying either wild-type or susceptibility alleles reversed (T > G) in a specific and concentration dependent manner with levels of wild-type (T) and risk (G) transcripts becoming equal under conditions of optimal inhibition with anti-miR-1324 (Fig. 4C). Susceptibility allele-specific loss of miR-1324 mediated silencing of INO80B in trophoblast cells. In human extravillous trophoblast cells (SGHPL5), a significant, about 2-fold increase of the INO80B chromatin-assembly complex transcript level was observed for the transcript expressing the risk allele (G) compared with the transcript expressing the wild-type allele (T) (**A**). After microRNA inhibition with anti-miR-409-5p, the levels of the INO80B transcript carrying the susceptibility allele (G) remained similar to the default situation (no inhibition) with significantly higher levels of the INO80B transcript carrying the wild-type allele (G) under all concentrations tested (**B**). In contrast, during inhibition with anti-miR-1324, the INO80B transcript levels carrying the wild-type or susceptibility alleles reversed (T > G) in a specific and concentration-dependent manner with levels of T and G alleles becoming equal under conditions of optimal inhibition (**C**). T: wild-type allele, G: risk allele; 20 000–50 000: cell number per well at start of transfection while other parameters such as DNA/Fugene HD ratio and total amount of DNA transfected were kept identical.

DISCUSSION
The present approach of pathway-guided genome-wide meta-analysis permits robust targeting of susceptibility genes that converge in a common pathway of multifactorial diseases with different founder effects. As demonstrated in the present study, this led to the rapid identification of the INO80B gene as cause of pre-eclampsia in Icelandic families. The implications are manifold. The 2p13 locus is also linked (LOD 3.5) with a subset of pre-eclamptic families in the Australian population (5). This association was seen for TACR1. Our data indicate that the adjacent gene, INO80B, is the cause of this association. Our approach could also allow identification of the additional loci for pre-

eclampsia in other founder populations (e.g. Finland, Norway) (24). To facilitate this, we also performed meta-analysis for the additional seven loci (Table 1) with demonstrated genome-wide linkage. These data sets are available upon request.

Secondly, we show that microRNA-mediated regulation of chromatin-remodeling complexes as previously demonstrated to regulate a distinctive step in the development of the vertebrate nervous system (12) appears essential for placental development as well. The use of conserved pathways in the brain and the placenta has been noted before (9) and has been confirmed recently (12). In the presence of the susceptibility allele (G), as expressed in the placental cells of children born from Icelandic pre-eclamptic females, seed-guided binding of miR-1324 to the ultraconserved 7-mer in the 3UTR is less efficient with upregulation (gain-of-function) of the cell cycle associated, chromatin-remodeling INO80B protein. MiR-1324 is restricted to Homininae (human, gorilla, orangutan, chimpanzee) which are the species with a hemochorial, discoid and villous placenta, deep invasion by extravillous trophoblast cells and extensive spiral artery remodeling (25,26).

The INO80B protein contains a novel winged helix domain (pfam14465) previously identified in INO80G (10). This domain (also known as NFRKB_WHL) has a significant structural similarity (Dali Z-score 6.7) (measure of similarity by comparing intramolecular distances) to one of the winged helix motifs within the C-terminal domain (CTD) of the cullin protein portion of the of the Cul1–Rbx1–Skp1–F boxSkp2 SCF ubiquitin ligase complex. This domain follows three repeats of the cullin repeat and is involved in binding of the RING finger protein Rbx1. NFRKB_WHL and the winged helix subdomain of CTD superpose with a RMSD of 2.15 Å over 63 residues and share a sequence identity of 10%. Thus, it is possible that NFRKB_WHL may be involved in protein–protein interactions rather than DNA binding as seen for STOX1 (10).

We do not exclude the involvement of additional factors in microRNA-mediated gene silencing of INO80B. The additional conserved sequences and the PRE-like element in the INO80B-3UTR are very interesting and indicate the involvement of additional proteins such as those of the Pumillio family. In fact, the necessity for Pumillio proteins is implemented in the MREdictor program as used in our study (15). The elaboration of this and other regulatory features (such as AGO2) however goes beyond the question presently addressed. Nevertheless, assuming that the protein involved in binding to the PRE-like element belongs to the Pumillio family, does not involve PUM1 or-2, and that this protein is differentially expressed in the pre-eclampsia/HELLP pathway, we checked our dataset for proteins that qualified for these criteria. Two Pumillio domain-containing candidates were found: NOP9 (C14orf21) and KIAA0020. This in combination with the other features we present will allow full, detailed and targeted exploration of the interactions involved allowing mechanistic insight.

In this respect, the functional picture that begins to emerge from the combined data (1,2, present study) is that the central, most consistent and thereby informative phenotypic defect in early-onset pre-eclampsia involves the essential and biologically central step in extravillous trophoblast differentiation when these cells terminally differentiate into giant cells (polyploidy). Premature differentiation such as by gene defects in STOX1, LINC-HELLP and possibly INO80B generates fetal cells that are halted too early in this process of differentiation and thereby in their normal function (maternal vessel wall adaptation). In humans, this is reflected, either preceded, accompanied or followed, by changes in cell cycle exit and/or defects in trophoblast invasion for the pre-eclampsia genes we tested. Experimentally, depending on the assay used, the outcome can involve both (cell cycle and invasion), only one (cell cycle or invasion) or none. At this stage, rather than applying an assay with less than superior qualities, we favor the establishment and implementation of a novel, better assay focusing on the above central feature (giant cell transition). The SGHPL5 cell line could permit such an assay. As a normal phenotypic feature reflecting function and in accordance with the fact that this and related cell lines are model systems for early extravillous trophoblast, these cells contain a low percentage of end-stage differentiated giant cells. If a protein or transcript can be identified with expression restricted to giant cells, a reliable quantitative parameter can be implemented that can be analyzed quantitatively and qualitatively in functional assays of extravillous trophoblast cells. Proteins of the REST/NRSF complex might be worth testing (Oudejans et al., unpublished). By this, an informative assay highly informative for the phenotype associated with pre-eclampsia can be established that retains the experimental versatility of the SGHPL5 and related model systems.

Clinically, for early diagnostic purposes and the design of truly effective therapeutic methods, future efforts should focus on non-invasive prenatal analysis of the placental transcriptome by genome-wide RNA sequencing methods using circulating RNA isolated from maternal plasma obtained during first trimester. Recent data show that this is technically feasible and highly informative (27,28). Given that the placental genes involved in early-onset pre-eclampsia and identified so far qualify as master control genes with an effect on multiple downstream genes (1,2), the RNA signatures in pregnancy-

associated diseases with placental origin will likely present in maternal plasma as signatures (i.e. specific gene sets of up- and downregulated genes) of a central theme (i.e. dysfunctional pathway). At the time analyzed (first trimester), the RNA signatures identified will not be obscured by the secondary additional differentially expressed RNA transcripts that generate additional signatures when the maternal system becomes activated and decompensated (second and third trimester). The information obtained by these genome-wide approaches will include the large layers of novel biological information (non-coding RNA inaccessible by proteomics or metabolomics), will allow presymptomatic diagnostics prior to the onset of maternal symptoms, might discriminate between pre-eclampsia subtypes and related syndromes and could be helpful for the design of novel, truly effective therapeutic interventions.

Finally, the combined data indicate a common theme underlying the genetic cause of pre-eclampsia and related syndromes. The placentally-expressed genes (STOX1, linc-HELLP, INO80B) involved in familial early-onset pre-eclampsia and the HELLP syndrome involve master control genes regulating multiple effector genes that control extravillous trophoblast invasion. The synteny of the chromosomal regions (10q22, 2p13) involved and the crosstalk or parallel function of these genes in the same pathway (trophoblast invasion) indicate a conserved fundamental mechanism (trophoblast cell cycle exit inversely related with invasion) that evolved as a consequence of hemochorial placentation. Genetically, this is reflected by different founder effects (Iceland, Netherlands, Australia) in paralogous chromosomal regions (2p13, 10q22) linked with the same disease (pre-eclampsia).

MATERIALS AND METHODS

Pathway-guided genome-wide meta-analysis

Transcripts (n = 5041) with significant (q-value 0.05) changes in quantitative or qualitative expression after siRNA-mediated knockdown of the HELLP lincRNA in extravillous trophoblast cells (SGHPL5) were selected for their location within the minimal critical region on 2p13 (D2S292-D2S329) linked with pre- eclampsia in Icelandic (lod 4.77) and Australian (lod 3.5) families (2,4,5). Out of the resulting set of 17 genes, 13 were prioritized for targeted exon sequencing.

Mutation analysis

Genomic DNA was isolated from EDTA blood of five informative affected females (IV:3, IV:4 from family A and III:2, III:3 and V:1 from family B) from the Iceland study and used for mutation analysis (4). Patients IV:3 and IV-4 are affected sisters. Patients III:2 and III:3 are cousins, their fathers are brothers. The grandmother of patient V:I and patients III:2 and III:3 are sisters. Coding exon specific primers were designed using ExonPrimer and used for mutation analysis as described (1,2). ExonPrimer was used with default settings including masking against hg18 (score 300) except that the GC clamp option was not used. Primer sequences are available upon request. Where appropriate, novel alternative exons as identified by the high-resolution RNA sequencing performed in extravillous trophoblast were included (2). Sequence files (.abl) were analyzed using 4Peaks, and aligned using Sequencher v5.

Haplotype analysis

Sequence variants identified were verified for their chromosomal position (UCSC hg19), screened for their class, function, molecule type, heterozygosity and allele frequencies (all SNPs: dbSNP 137) and scored for maximal allele sharing between affected females. Variants were considered mutations when absent in 100 normal controls (200 chromosomes).

High-density genome-wide scan

The Affymetrix CytoScan HD (Affymetrix, catalog no. 901835) contains 750 000 unique SNPs and 1.9 million oligonucleotide probes with a mean backbone spacing of one oligonucleotide probe every 2 kb and one oligonucleotide probe every 400 basepairs in targeted regions. The mean number of SNP probes is 200 per megabase. All probes are 25 bp long. Each SNP is targeted by six probes, three for each allele. Approximately, 90% of SNPs have a minor allele frequency of 0.05 or greater. The high-density (HD) assay was run according to the manufacturer's protocol (Affymetrix). In brief, 250 ng of patient DNA was digested with NspI, amplified with Titanium Taq DNA polymerase, fragmented with Affymetrix fragmentation reagent and labeled with biotin-end labeled nucleotides. The DNA was hybridized to the microarray for 16 h, washed on the GeneChip Fluidics station 450, stained with Affymetrix GeneChip Stain reagents and scanned on the GeneChip Scanner 3000 7G. Data analysis was performed using Nexus Copy Number Software version 6.1.

RNA expression analysis

Strand-specific RNA analysis was done as described (1,2).

MicroRNA expression quantification

Total RNA was isolated from SGHPL5 cells (resuspended in RNAbee) by Qiagen affinity column isolation with on-column DNase digestion. First strand synthesis was done with 10 ng total RNA with specific RT primers using MultiScribe Reverse Transcriptase (Life Technologies) for 30 min at 16°C, 30 min at 42°C and 5 min at 85°C in a volume of 5 µl. For pre-amplification, 2.5 µl of Taqman pre-amplification master mix (Life Technologies), 1.25 µl of 5-fold diluted assay mix (Life Technologies) and 1.25 µl of RT-reaction mix were combined, heated for 10 min at 95°C, followed by 14 cycles for 15 s at 95°C and 4 min at 60°C. For quantification, 4.5 µl of pre-amplification mix was combined with Taqman assay and universal master mix, heated for 2 min at 50°C, for 10 min at 95°C followed by 40 cycles for 15 s at 95°C and 1 min at 60°C using an ABI 7300 PCR machine. MicroRNAs analyzed were miR-24 (Life Technologies 000402) and miR-1324 (Life Technologies 002815). All reactions were done in duplicate. Negative controls consisted of identical reactions with omission of the reverse transcriptase enzyme. Data analysis was done with SDS (version 2.3) and RQmanager (version 1.2).

MicroRNA-target interaction analysis

A conserved 70 bp fragment of the 3UTR of INO80B containing the wild-type (T) or susceptibility (G) allele of rs34174194 was obtained by PCR using genomic DNA from homozygous samples and directionally cloned into SpeI and HindIII sites of pMIR-report luciferase vector (Life Technologies). The pMIR-REPORT™ Luciferase miRNA Expression Reporter Vector contains firefly luciferase under the control of a mammalian promoter/terminator system and with an miRNA target cloning region downstream of the luciferase translation sequence. The pMIR-REPORT-INO80B-3UTR construct with the wild-type allele (T) of rs34174194 contained the sequence: 5′-ACT AGT GTG GGT CGA ACC TTC ACC TAC TCT TTG TCT GTA CAA GCT CTT G**T**T GTC TGG GTG AGC GGC GGA GGC GCT GAA GCTT-3′. The pMIR-REPORT-INO80B-3UTR construct with the susceptibility allele (G) of

rs34174194 contained the sequence: 5′-ACT AGT GTG GGT CGA ACC TTC ACC TAC TCT TTG TCT GTA CAA GCT CTT G**G**T GTC TGG GTG AGC GGC GGA GGC GCT GAA GCTT-3′. Inserts were verified by DNA sequencing. Plasmids were calibrated by quantitative PCR (Supplementary Material, File S10) to assure transfection with equimolar concentrations of wild-type and susceptibility allele constructs with the only difference being restricted to the variant nucleotide (T or G). SGHPL5 cells (human extravillous trophoblast) were cultured in Iscove's Modified Dulbecco's Medium (IMDM) with Glutamax, penicillin/streptomycin and Fetal Bovine Serum (10% v/v) (heat inactivated) (Life Technologies). At Day 1 of transfection, near-confluent cells (80%) were harvested by trypsinization, washed, resuspended and counted using the Countess Automatic Cell Counter (Invitrogen). Cells were plated at a concentration of 10 000–50 000 cells in a volume of 200 µl complete medium per well in 48 well polystyrene plates. At Day 2, medium was replaced and cells were co-transfected (in triplicate) with previously determined optimal (highest signal, lowest toxicity) concentrations (50 ng) of pMIR-reporter constructs, either wild-type (T) or mutant (G), together with equal concentrations of a second vector, pMIR-REPORT™ Beta-galactosidase Reporter Control Vector for normalizing transfection efficiency. Transfection was done with Fugene HD using a 3:1 ratio of Fugene HD/DNA. DNA/Fugene HD complexes were prepared in serum-free medium according to the manufacturers's instructions (Promega). For competition with microRNA inhibitors, all experiments were repeated as above except that transfection mixes included 50 nM of mirVana microRNA inhibitors against hsa-miR-574-3p (MH12848), hsa-miR-409-5p (MH13028) or hsa-mIR-1324 (MH13340) (Life Technologies). miRNA inhibitors are small, chemically modified single-stranded RNA molecules designed to specifically bind to and inhibit endogenous miRNA molecules and enable miRNA functional analysis by down-regulation of miRNA activity. For the siRNA inhibitor experiments, the recommendations of the manufacturer were followed. For siRNA experiments to be valid and as based on the one-tailed t-test, at a fold-difference of 2 for the biological effect tested (T versus G allele) and with a %CV of 21–25% (as seen in our assays) at least four replicate observations need to be obtained with each observation measured in triplicate. To prevent the problem of pseudoreplication (29), independent measurements separated in time are usually considered essential for hierarchical systems (such as cell cultures), i.e. the whole procedure repeated three times on different days, weeks or months. However, these measurements still remain only technical replicates, while becoming subject to inter-assay variability. We therefore used a different approach. We introduced of a dose-dependent variable parameter directly related to the biological question addressed (i.e. creating a higher order hierarchical system). For this, within the previously determined optimal window of cell density, a dose response effect was created by the use of 4 different cell concentrations (20 000–50 000), while keeping the amount of transfected DNA and siRNA at constant, but previously determined optimal and non-toxic levels. This prevents the need for measurements separated in time (ruling out inter-assay variability), while the dose response curve generates additional control criteria that have to be met in the readout, in dosage effect (increase) and/or directional change (decrease after inhibition). The statistical analysis needed in this setup needs to be adapted likewise and was applied (two-way ANOVA). Luciferase and galactosidase measurements were done 48 h after transfection according to the manufacturer's instructions (Promega) with background corrections with the inclusion of untransfected cell lysates. Signals were considered valid when levels reached at least three times the background signal. Galactosidase was measured at 405 nm using the Synergy HT Multi-mode microplate reader (Gen5 software) (BioTek) and only considered valid if values were at least 0.1. Luciferase was measured using a Lumat LB 9507 Luminometer (Berthold Technologies). Activity was calculated as Relative Light Units (RLU) (luciferase/BETA-gal) and analyzed using GraphPad Prism version 6.0c. Statistical significance was determined by two-way ANOVA.

URLS
SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

REFERENCES

1   van Dijk M., Mulders J., Poutsma A., Könst A.A., Lachmeijer A.M., Dekker G.A., Blankenstein
        M.A., Oudejans C.B. Maternal segregation of the Dutch preeclampsia locus at 10q22 with a
        new member of the winged helix gene family. Nat. Genet. 2005;37:514-519
2   van Dijk M., Thulluru H.K., Mulders J., Michel O.J., Poutsma A., Windhorst S., Kleiverda G.,
        Sie D., Lachmeijer A.M., Oudejans C.B. HELLP babies link a novel lincRNA to the trophoblast
        cell cycle. J. Clin. Invest. 2012;122:4003-4011
3   Zintzaras E., Kitsios G., Harrison G.A., Laivuori H., Kivinen K., Kere J., Messinis I., Stefanidis
        I., Ioannidis J.P. Heterogeneity-based genome search meta-analysis for preeclampsia. Hum.
        Genet. 2006;120:360-370
4   Arngrímsson R., Sigurõardóttir S., Frigge M.L., Bjarnadóttir R.I., Jónsson T., Stefánsson H.,
        Baldursdóttir A., Einarsdóttir A.S., Palsson B., Snorradóttir S., et al. A genome-wide scan
        reveals a maternal susceptibility locus for pre-eclampsia on chromosome 2p13. Hum. Mol.
        Genet. 1999;8:1799-1805. Abstract/FREE Full Text
5   Fitzpatrick E., Göring H.H., Liu H., Borg A., Forrest S., Cooper D.W., Brennecke S.P., Moses
        E.K. Fine mapping and SNP analysis of positional candidates at the preeclampsia
        susceptibility locus (PREG1) on chromosome 2. Hum. Biol. 2004;76:849-862   ↵
        Doridot L., Passet B., Méhats C., Rigourd V., Barbaux S., Ducat A., Mondon F., Vilotte M.,
        Castille J., Breuiller-Fouché M., et al. Preeclampsia-like symptoms induced in mice by
        fetoplacental expression of STOX1 are reversed by aspirin treatment. Hypertension
        2013;61:662-668
6   Chambers A.L., Ormerod G., Durley S.C., Sing T.L., Brown G.W., Kent N.A., Downs J.A. The
        INO80 chromatin remodelling complex prevents polyploidy and maintains normal chromatin
        structure at centromeres. Genes Dev. 2012;26:2590-2603   ↵
7   Park E.J., Hur S.K., Lee H.S., Lee S.A., Kwon J. The human Ino80 binds to microtubule via
        the E-hook of tubulin: implications for the role in spindle assembly. Biochem. Biophys. Res.
        Commun. 2011;416:416-420.
8   van Dijk M., van Bezu J., Poutsma A., Veerhuis R., Rozemuller A.J., Scheper W.,
        Blankenstein M.A., Oudejans C.B. The pre-eclampsia gene STOX1 controls a conserved
        pathway in placenta and brain upregulated in late-onset Alzheimer's disease. J. Alzheimers
        Dis. 2010;19:673-679.
9   Kumar A., Möcklinghoff S., Yumoto F., Jaroszewski L., Farr C.L., Grzechnik A., Nguyen P.,
        Weichenberger C.X., Chiu H.J., Klock H.E., et al. Structure of a novel winged-helix like
        domain from human NFRKB protein. PLoS One 2012;7:e43761.
10   Grossman S.R., Andersen K.G., Shlyakhter I., Tabrizi S., Winnicki S., Yen A., Park D.J.,
        Griesemer D., Karlsson E.K., Wong S.H., et al. Identifying recent adaptations in large-scale
        genomic data. Cell 2013;152:703-713.
11   Buhimschi I.A., Nayeri U.A., Zhao G., Shook L.L., Pensalfini A., Funai E.F., Bernstein I.M.,
        Glabe C.G., Buhimschi C.S. Protein misfolding, congophilia, oligomerization, and defective
        amyloid processing in preeclampsia. Sci. Transl. Med. 2014;6:245-ra92.
12   Trapnell C., Williams B.A., Pertea G., Mortazavi A., Kwan G., van Baren M.J., Salzberg S.L.,
        Wold B.J., Pachter L. Transcript assembly and quantification by RNA-Seq reveals
        unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol.
        2010;28:511-515.
13   Yoo A.S., Staahl B.T., Chen L., Crabtree G.R. MicroRNA-mediated switching of chromatin-
        remodelling complexes in neural development. Nature 2009;460:642-646.
14   Incarnato D., Neri F., Diamanti D., Oliviero S. MREdictor: a two-step dynamic interaction
        model that accounts for mRNA accessibility and Pumilio binding accurately predicts
        microRNA targets. Nucleic Acids Res. 2013;41:8421-8433. Abstract/FREE Full Text
15   Kertesz M., Iovino N., Unnerstall U., Gaul U., Segal E. The role of site accessibility in
        microRNA target recognition. Nat. Genet. 2007;39:1278-1284.
16   Lewis B.P., Burge C.B., Bartel D.P. Conserved seed pairing, often flanked by adenosines,
        indicates that thousands of human genes are microRNA targets. Cell 2005;120:15-20.
17   Bartel D.P. MicroRNAs: target recognition and regulatory functions. Cell 2009;136:215-233.
18   Nam J.-W., Rissland O.S., Koppstein D., Abreu-Goodger C., Jan C.H., Agarwal V., Yildirim
        M.A., Rodriguez A., Bartel D.P. Global analyses of the effect of different cellular contexts on
        microRNA targeting. Mol. Cell 2014;53:1031-1041.
19   Luo S.S., Ishibashi O., Ishikawa G., Ishikawa T., Katayama A., Mishima T., Takizawa T.,

Shigihara T., Goto T., Izumi A., et al. Human villous trophoblasts express and secrete placenta-specific microRNAs into maternal circulation via exosomes. Biol. Reprod. 2009;81:717-729. Abstract/FREE Full Text

20 Morales-Prieto D.M., Chaiwangyen W., Ospina-Prieto S., Schneider U., Herrmann J., Gruhn B., Markert U.R. MicroRNA expression profiles of trophoblastic cells. Placenta 2012;33:725-734.

21 Afanasyeva E.A., Hotz-Wagenblatt A., Glatting K-H., Westermann F. New miRNAs cloned from neuroblastoma. BMC Genomics 2008;9:52.

22 Witten D., Tibshirani R., Gu S.G., Fire A., Lui W-O. Ultra-high throughput sequencing-based small RNA discovery and discrete statistical biomarker analysis in a collection of cervical tumors and matched controls. BMC Biol. 2010;8:58.

23 Laivuori H., Lahermo P., Ollikainen V., Widen E., Häivä-Mällinen L., Sundström H., Laitinen T., Kaaja R., Ylikorkala O., Kere J. Susceptibility loci for preeclampsia on chromosomes 2p25 and 9p13 in Finnish families. Am. J. Hum. Genet. 2003;72:168-177.

24 Brameier M. Genome-wide comparative analysis of microRNAs in three non-human primates. BMC Res. Notes 2010;3:64.

25 Crosley E.J., Elliot M.G., Christians J.K., Crespi B.J. Placental invasion, preeclampsia risk and adaptive molecular evolution at the origin of the great apes: evidence from genome-wide analyses. Placenta 2013;34:127-132.

26 Tsui N.B., Jiang P., Wong Y.F., Leung T.Y., Chan K.C., Chiu R.W., Sun H., Lo Y.M. Maternal plasma RNA sequencing for genome-wide transcriptomic profiling and identification of pregnancy-associated transcripts. Clin. Chem. 2014;60:954-962. Abstract/FREE Full Text

27 Koh W., Pan W., Gawad C., Fan H.C., Kerchner G.A., Wyss-Coray T., Blumenfeld Y.J., El-Sayed Y.Y., Quake S.R. Noninvasive in vivo monitoring of tissue-specific global gene expression in humans. Proc. Natl Acad. Sci. USA 2014;111:7361-7366. Abstract/FREE Full Text

28 Lazic S.E. The problem of pseudoreplication in neuroscientific studies: is it affecting yoanalysis? BMC Neurosci. 2010;11:5.