Synopsis of Preterm Birth Genetic Association Studies: The Preterm Birth Genetics Knowledge Base (PTBGene)

S.M. Dolan[a, o]   M.V. Hollegaard[b, o]   M. Merialdi[c, o]   A.P. Betran[c, o]   T. Allen[d]
C. Abelow[a]   J. Nace[e, o]   B.K. Lin[f]   M.J. Khoury[g]   J.P.A. Ioannidis[h, i]   S. Bagade[j]
X. Zheng[k]   R.A. Dubin[k]   L. Bertram[j, l]   D.R. Velez Edwards[m, o]   R. Menon[c, e, n, o]

[a]Department of Obstetrics and Gynecology and Women's Health, Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, N.Y., USA; [b]Section of Neonatal Screening and Hormones, Statens Serum Institute, Copenhagen, Denmark; [c]Department of Reproductive Health and Research and [d]Department of Knowledge, Management and Sharing, World Health Organization, Geneva, Switzerland; [e]The Perinatal Research Center, Nashville, Tenn., [f]Office of the Medical Director, March of Dimes, White Plains, N.Y., [g]National Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Ga., USA; [h]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, and Biomedical Research Institute, Foundation for Research and Technology – Hellas, Ioannina, Greece; [i]Center for Genetic Epidemiology and Modeling, Tufts Medical Center, and Department of Medicine, Tufts University School of Medicine, Boston, Mass., [j]Genetics and Aging Research Unit, MassGeneral Institute for Neurodegenerative Disease (MIND), Department of Neurology, Massachusetts General Hospital, Charlestown, Mass., [k]Bioinformatics Shared Resource (BISR), Albert Einstein College of Medicine, Bronx, N.Y., USA; [l]Neuropsychiatric Genetics Group, Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany; [m]Vanderbilt Epidemiology Center, Institute of Medicine and Public Health, Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, Tenn., [n]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Ga., [o]PREBIC Genetic Working Group 2007, USA

Abstract

Aim: Our goal was to produce a field synopsis of genetic associations with preterm birth and to set up a publicly available online database summarizing the data. *Methods:* We performed a systematic review and meta-analyses to identify genetic associations with preterm birth. We have set up a publicly available online database of genetic association data on preterm birth called PTBGene (http://ric.einstein.yu. edu/ptbgene/index.html) and report on a structured synopsis thereof as of December 1, 2008. *Results:* Data on 189 polymorphisms in 84 genes have been included and 36 meta-analyses have been performed. Five gene variants (4 in maternal DNA, one in newborn DNA) have shown nominally significant associations, but all have weak epidemiological credibility. *Conclusion:* After publishing this field synopsis,

The views, findings and conclusions expressed in this report are solely those of the authors and do not necessarily reflect the views, decisions or stated policy of the United States Department of Health and Human Services or the World Health Organization.

Siobhan M. Dolan, MD, MPH
Department of Obstetrics and Gynecology and Women's Health
Albert Einstein College of Medicine
Belfer 501, 1300 Morris Park Avenue, Bronx, NY 10461 (USA)
Tel. +1 718 430 4275, Fax +1 718 430 8774, E-Mail siobhanmdolan@yahoo.com

the PTBGene database will be regularly updated to keep track of the evolving evidence base of genetic factors in preterm birth with the goal of promoting knowledge sharing and multicenter collaboration among preterm birth research groups.

Introduction

Preterm birth is an international major public health concern, with rates high and rising in many parts of the world. In 2006, 12.8% of all births in the US were preterm, defined as delivery before 37 completed weeks of gestation [1]. Preterm birth is the second leading cause of infant mortality and the leading cause of mortality among black infants in the US as well as the major contributor to worldwide infant mortality and morbidity [2]. Children born preterm may suffer lifelong morbidities including lung disease, vision and hearing deficits and other neurosensory impairments [3] as well as predisposition to hypertension and diabetes in adult life [4]. While many environmental contributors to preterm birth have been proposed (e.g. infection, stress and smoking), most preterm births remain unexplained. Several pathways leading to preterm birth have been outlined, including (1) inflammation, (2) maternal/fetal stress, (3) uterine distension and cervical dysfunction, (4) decidual hemorrhage, and (5) metabolic enzyme variation [5, 6]. A modest body of literature has examined the complex genetic contribution to preterm birth [7–9], with heritability estimates in the range of 27–36% [10, 11]. Several studies have reported associations with single gene variants [12–21]. These studies have been hampered by varying definitions of preterm birth, as well as small sample sizes and failure of replication, possibly reflecting publication bias. It remains a challenge to identify robust associations between genetic variation and preterm birth risk.

Material and Methods

In conjunction with the Human Genome Epidemiology Network (HuGENet), the World Health Organization, Albert Einstein College of Medicine, Massachusetts General Hospital, and March of Dimes, members of the Preterm Birth International Collaborative (PREBIC) have created PTBGene, a regularly updated online, publicly available database of all genetic association studies published in the field of preterm birth to date. This is modeled after an approach to creating field synopses for Alzheimer's disease and schizophrenia developed at Massachusetts General Hospital [22, 23] and following the recommendations for reporting proposed by the meta-analysis of observational studies in epidemiology group [24]. PTBGene is implemented using Oracle database technology. Data are captured in a relational database and the web application was developed using the PL/SQL web toolkit. PTBGene is accessible at http://ric.einstein.yu.edu/ptbgene/index.html.

Literature Searches

This research was reviewed by the Institutional Review Board at Montefiore Medical Center and was found to be exempt. MEDLINE and EMBASE were searched using a comprehensive search strategy (table 1) to identify all studies reporting on genetic associations with preterm birth in humans published between January 1, 1990 and December 1, 2008, with no language restrictions. The search strategy was designed to be broad and inclusive since preterm birth can be referred to in many ways, including preterm birth, preterm delivery, preterm labor, premature birth, etc. The phenotype of interest was defined as singleton preterm birth. The search identified 6,772 abstracts. Eight authors are PREBIC members (S.M.D., M.V.H., M.M., A.P.B., J.N., B.K.L., D.R.V., R.M.). Based on their expertise in preterm birth, they screened the abstracts, reviewed the articles and extracted the data in pairs. The abstracts were screened independently by 2 authors according to the following inclusion and exclusion criteria to identify the studies, providing the pertinent information regarding singleton preterm birth in humans. They had to provide the case definition of preterm birth in gestational weeks at delivery and genotypes for the cases and controls (maternal, newborn and/or paternal). Studies exclusively of twins were excluded, as were case only studies, case reports, family-based studies, and review articles.

The screening yielded 139 abstracts for which the full-text articles were obtained and archived electronically and in paper version. Two different authors then reviewed the articles, compared their independent selections and came to a consensus on the 61 articles that met the inclusion and exclusion criteria. These 61 articles had data extracted (table 2).

To assess our ability to identify the appropriate articles, we compared our results to those of HuGE Navigator [25], an integrated, searchable knowledge base of genetic associations and human genome epidemiology [26], and the Genetic Association Database (GAD) [27] which were searched for preterm birth. HuGE Navigator identified 67 articles and GAD identified 68, but these included review articles and articles examining pediatric outcomes after preterm birth such as retinopathy of prematurity, which we excluded. We checked to confirm that we had considered all articles found by HuGE Navigator and GAD.

Data Extraction

The data extraction was performed by one author (PREBIC member) and checked by a second. A set of data for each variant was extracted from the articles and entered into a spreadsheet, including the gene name and variant as identified by the official symbol and GeneID in NCBI's Entrez Gene [28], and the rs number from dbSNP [29]. The first author's last name, year and journal of publication and PubMed ID were extracted. The geographic location of the population studied was recorded as city and country. One field identified whether the variant was reported in maternal DNA, newborn DNA or paternal DNA. The definition of preterm in gestational weeks was recorded as was the case definition of preterm, categorized as (1) spontaneous preterm birth

Table 1. Search strategy for use in MEDLINE and EMBASE

Databases: EMBASE + MEDLINE
Limited to human
Years: January 1, 1990 to December 1, 2008

Step 1
'immaturity' or 'premature labor' or prematurity or 'premature birth*' or prematurity or 'low birth weight' or 'low-birth-weight infant' or 'very low birth weight' or 'very-low-birth-weight infant' or 'premature labour' or 'premature labor' or 'preterm birth' or 'pre-term infant*' or 'preterm infant*' or 'pre-term bab*' or 'preterm bab*' or 'pre-term deliver*' or 'preterm deliver*' or 'premature deliver*' or 'pre-mature deliver*' or 'pre-term birth*' or 'preterm birth*' or 'premature bab*' or 'pre-mature bab*' or 'premature infant*' or 'pre-mature infant*' or 'premature infant' or 'premature infant*' or 'low birth weight' or 'very low birth weight' or 'small for date infant' or 'small for gestational age'

Step 2
'pregnancy complication*' OR 'pregnancy complication' OR 'pregnancy' or 'pregnan*' or 'gestational age'

Step 3
'premat*' or 'pre-matur*' or 'early labour*' or 'early labor*' or 'preterm*' or 'pre-term*' (Step 2 AND Step 3) or Step 1

Step 4
'population genetic parameters' or 'environmental carrying capacity' or 'familial incidence' or 'gene frequency' or 'genetic distance' or 'genetic drift' or 'genetic stability' or 'genetic variability' or 'microsatellite instability' or 'neutral gene theory' or 'combining ability' or 'genetic linkage' or 'heredity' or 'allelism' or 'cancer genetics' or 'genealogy' or 'evolutionary adaptation' or 'genetic compatibility' or 'genetic cross' or 'genetic damage' or 'genetic heterogeneity' or 'genetic line' or 'genetic resource' or 'genome' or 'genotype' or 'human genetics' or 'hybridization' or 'immunogenetics' or 'inheritance' or 'autosomal inheritance' or 'dominant inheritance' or 'multifactorial inheritance' or 'genetic predisposition' or 'recessive inheritance' or 'sex chromosomal inheritance' or 'zygosity' or 'mathematical genetics' or 'mutation' or 'pharmacogenetics' or 'phenotype' or 'population genetics' or 'quantitative genetics' or 'genetic and familial disorders' or 'familial disease' or 'genetic disorder' or 'chromosome aberration' or 'hereditary connective tissue disease' or 'hereditary optic atrophy' or 'inborn error of metabolism' or 'monogenic disorder' or 'multifactorial genetic disorder' or 'genetic*'

Step 5
Step 4 AND ([Step 2 AND Step 3] or Step 1)

6,772 citations identified.

(defined as preterm birth following spontaneous preterm labor), (2) preterm birth following preterm premature rupture of the membranes (PPROM), (3) preterm birth following maternal medical complications such as preeclampsia and diabetes, (4) preterm birth following fetal complications such as small for gestational age or intrauterine growth restriction, or (5) preterm birth not

Table 2. Systematic review

	Number of articles	Number of genes	Number of polymorphisms
Overall data set	61	84	189
Maternal data	46	48	127
Newborn data	31	68	145
Paternal data	2	3	4

6,772 abstracts were identified and screened. 139 full-text articles were retrieved. 61 articles had data extracted.

otherwise specified. In some cases, 2 codes were assigned for case definition if, for example, the cases reported included spontaneous preterm births as well as those following PPROM. Also, the reported ancestry of the mothers in the study was coded according to a scheme used previously [22, 30]. The 5 general groups of origin were (1) African, (2) Asian, (3) Caucasian (of European descent), (4) Hispanic, and (5) mixed/other. For each polymorphism (usually rs ID) and ethnic group, the number of cases and controls was recorded as well as the major and minor allele for the single nucleotide polymorphism (SNP). In cases where there was inconsistency in reporting major versus minor allele between studies, we delineated the ancestral allele at dbSNP [29] as the major allele. Additionally, the number of cases and controls with each genotype (minor/minor, minor/major, major/major) was recorded. In cases where genotypes were not provided in the articles, the corresponding authors were contacted and asked to for the genotype data which were then included when provided. Only corresponding authors who were also PREBIC members responded to our queries, but as more researchers learn of PTBGene, we expect that more extensive data can be captured. This constituted the data set for the meta-analyses.

Statistical Analysis
Whenever there were 3 or more different datasets evaluating the same genetic variant in mothers, fathers or newborns, a meta-analysis was performed. The primary analyses combined data on all relevant studies and a secondary analysis performed separate analyses according to racial descent.

We used the odds ratio (OR) as the metric of choice for all meta-analyses. In the absence of strong biological rationale for any specific genetic model for any of the probed associations, the main analyses were performed using allele-based comparisons. Calculations were performed with Statistical Analysis Software (SAS), version 9.1.3. All p values are two-sided.

Statistical heterogeneity among the studies was evaluated by Cochran's Q [31], considered significant at $p < 0.10$. Both fixed-effects and random-effects models were used for summary effects. However, since the Q test is insensitive to small/few studies, we based our main inferences on the random-effects model. This model assumes that the studies are a random sample of a hypothetical population of studies taking into account within- and between-study variability. We also used the I^2 metric [31, 32] as a measure of the extent of between-study heterogeneity; values of 50% or higher are considered to express large between-study het-

erogeneity and values of 25–50% express moderate between-study heterogeneity. With a small number of studies, I^2 can have large uncertainty, so inferences should be interpreted cautiously [32].

We performed sensitivity analyses exploring the possibility for bias. For the formally statistically significant associations (those with $p < 0.05$ on random-effects calculations) where an adequate number of studies existed, we evaluated whether the results were different after exclusion of the first-published study and after exclusion of studies that had significant ($p < 0.05$) violation of Hardy-Weinberg equilibrium (HWE) in the controls, according to an exact test [33]. We did not apply regression diagnostics for asymmetry that may reflect publication bias, other biases, quality differences, or genuine heterogeneity between small and larger studies, since all meta-analyses had a limited number of studies [34, 35]. For the same reason, we did not apply tests for excess of single studies with significant findings [36].

Assessment of Cumulative Evidence

We applied a grading system to each nominally significant association that assesses the epidemiological strength of the cumulative evidence through a process in which the first author (S.M.D.) assigned the grades which were then checked by the epidemiologist who presented the criteria in detail elsewhere (J.P.I.) [37]. There were no disagreements. Briefly, each meta-analyzed association is graded based on the amount of evidence, replication and protection from bias in an objective fashion according to the following guidelines. For amount of evidence, grade A is assigned when the total number of minor alleles of cases and controls combined is >1,000, B when it is 100–1,000 and C when it is <100. For replication consistency, C reflects point estimates of $I^2 > 50\%$, B is 25–50% and A is <25%. For protection from bias, grade A means that bias, if present, may change the magnitude but not the presence of an association; grade B means that there is no evidence of bias that would invalidate an association, but important information is missing; and grade C means that there is demonstrable potential or clear bias that may invalidate the mere presence of an association. We consider various potential sources of bias by carefully reviewing methods, including errors in phenotypes, genotypes, confounding (population stratification), and errors/biases at the meta-analysis level (publication and other selection biases). When the summary OR deviates less than 1.15-fold from the null, for meta-analyses based on published data, we acknowledge that occult publication and selective reporting biases alone may invalidate the association, regardless of the presence or not of other biases, therefore the grade is C. When the summary OR deviates more than 1.15-fold from the null, grade C is given when nominal statistical significance is lost with the exclusion of the first-published study or with correction for HWE-violating studies. For large OR estimates, we consider that phenotype misclassification may affect the magnitude but not the presence of an effect in this field; genotyping errors are also considered to affect the magnitude but usually not the presence of significant associations of such magnitude. Potential confounding from population stratification is also considered to have a similar impact (given that at least self-reported racial descent is taken into account in all our analyses).

Associations that get 3 A grades are considered to have 'strong' epidemiological credibility. Associations that get any B, but not any C grade, are assigned 'moderate' credibility, and associations that get any C grade are considered to have 'weak' credibility.

Results

Sixty-one eligible articles provided data on 189 polymorphisms in 84 genes. These included 46 articles reporting data on maternal DNA (127 polymorphisms in 48 genes), 31 articles reporting newborn DNA (145 polymorphisms in 68 genes) and 2 articles reporting paternal DNA (4 polymorphisms in 3 genes).

Of the 46 articles that reported on maternal genotypes, 32 defined preterm birth as birth before 37 completed weeks of gestation, 6 articles used a definition of 36 weeks, 4 used 35 weeks, 2 used 34 weeks, and 2 used 32 weeks. Of the 31 articles reporting on newborn genotypes, 15 used a 37 week cut-off, 6 used 36 weeks, 7 used 35 weeks, one 34 weeks, one 30 weeks, and one 27 weeks. The 2 articles that reported paternal genotypes used 37 weeks as the case definition.

In the maternal reports, 23 articles defined a case as spontaneous preterm birth, 1 defined it as following PPROM, 8 defined it as spontaneous preterm birth or preterm birth following PPROM, and 14 did not define the case beyond preterm. In the newborn reports, 10 articles defined a case as spontaneous preterm birth, 1 defined it as following PPROM, 3 combined spontaneous and PPROM cases, 1 combined maternal and fetal indications as well as spontaneous preterm birth in the case definition, and 16 did not define the case beyond preterm. Neither of the 2 paternal reports had a detailed case definition.

Overall only 19 of the 46 maternal and 14 of the 31 newborn studies had more than 100 cases. Sufficient data to perform meta-analysis were available for 22 polymorphisms in 15 genes for maternal DNA and 14 polymorphisms in 8 genes for newborn DNA (reflecting 107 and 62 datasets, respectively). No sufficient data were available to examine paternal genetic associations. We did not combine paternal with maternal data for analysis as the paternal contribution to preterm birth is thought to be much smaller than the maternal [38].

Of the 22 polymorphisms in 15 genes studied in meta-analyses for maternal DNA variants, 4 variants, including an 86 bp VNTR in interleukin 1 receptor antagonist (*IL1RN*) and 3 SNPs in beta-2 adrenergic receptor (*ADRB2*, rs1042713), interferon gamma (*IFNG*, rs2430561), and factor 2 (*F2*, rs1799963) showed nominally significant summary ORs with modest p values in the main (allele-based) analyses (ORs ranging from ~1.4 to ~1.8, p values ranging from 0.01 to 0.04, table 3). Of the 14 polymorphisms in 8 genes studied in meta-analyses of newborn DNA variants, the same SNP in coagula-

Table 3. Random-effects meta-analysis using allelic contrasts for polymorphisms showing significant summary ORs (as of December 1, 2008)

Gene	Polymorphism	Alleles	OR (95% CI), p value	Hetero-geneity	Allele count in minor genetic group	Datasets
IL1RN (maternal)	86 bp VNTR	2 vs. 1 copy	1.77 (1.11–2.81) p = 0.02	$I^2 = 36$ p = 0.18	211	5
ADRB2 (maternal)	rs1042713	A vs. G	0.60 (0.41–0.88) p = 0.01	$I^2 = 28$ p = 0.25	516	3
IFNG (maternal)	rs2430561	A vs. T	1.43 (1.08–1.91) p = 0.01	$I^2 = 0$ p = 0.71	508	4
F2 (maternal)	rs1799963	A vs. G	1.84 (1.04–3.28) p = 0.04	$I^2 = 0$ p = 0.55	51	4
F2 (newborn)	rs1799963	A vs. G	1.81 (1.05–3.12) p = 0.03	$I^2 = 15$ p = 0.32	75	5

tion factor 2, also showing association in the maternal analyses (*F2*, rs1799963), yielded a nominally significant OR of 1.81 (p = 0.03) (table 3). The 5 respective forest plots are shown in figure 1.

Functional Considerations for Identified Nominally Significant Associations

Little is known about the precise mechanisms by which the identified genes contribute to preterm birth. Brief functional considerations are provided, but further research is warranted to learn about mechanisms of action.

The *IL1RN* gene (GeneID 3557) [28] encodes a protein that inhibits the activities of *IL1A* (interleukin 1 alpha) and (*IL1B*) and interleukin 1 beta. The *IL1* cytokine family is involved in immune and inflammatory responses and the 86 base pair VNTR in *IL1RN* may therefore play a role in preterm birth through its effect on inflammatory response.

ADRB2 is the gene that encodes beta-2-adrenergic receptor, a member of the G protein-coupled receptor superfamily (GeneID 154) [28]. The nominally associated coding exon SNP rs1042713 results in the amino acid change Arg16Gly which is hypothesized to enhance the rate of β-agonist-induced desensitization and thus to protect against preterm birth [39, 40].

IFNG encodes interferon gamma, GeneID 3458 [28]. *IFNG* is involved in transcriptional regulation of a variety of genes involved in the immune response and coordinates a diverse array of cellular programs [41]. The nominally associated rs2430561 correlates with lower levels of interferon gamma production which may mitigate the immunotolerance required for pregnancy maintenance and increase risk for preterm birth [20, 42].

The last nominal association is with rs1799963 in *F2*, coagulation factor II (GeneID 2147) [28]. Coagulation factor II is involved in the first step of the coagulation cascade in which it is proteolytically cleaved to form thrombin. *F2* variants can be associated with thrombosis, and the fact that *F2* was identified in both maternal and newborn samples suggests that thrombosis may play a key role in the etiology of preterm birth.

Grading the Evidence

None of the 5 nominal associations had formal statistical significance when analyses were limited to studies of Caucasian descent populations, but effect sizes did not change markedly. Using interim guidelines to assess the 'epidemiological credibility' of these associations [37], we found that for 'amount of evidence', 3 nominal associations in maternal DNA were graded B, while the associations for rs1799963 in the maternal and newborn analyses were graded C. I^2 was less than 25% in 3 of the 5 associations (grade A) and 36% and 28% in the other 2 (grade B). The magnitude of the effect deviated more than 1.15-fold from the null for all 5 associations, but 3 lost their nominal significance and the other 2 had less than 3 studies left when the first study was excluded. Moreover, the phenotype definitions left room for considerable bias. Therefore, all 5 associations were graded C for protection from bias. As a result, all 5 nominal associations were graded as having only 'weak' epidemiological credibility (overall grade C) per interim guidelines (table 4).
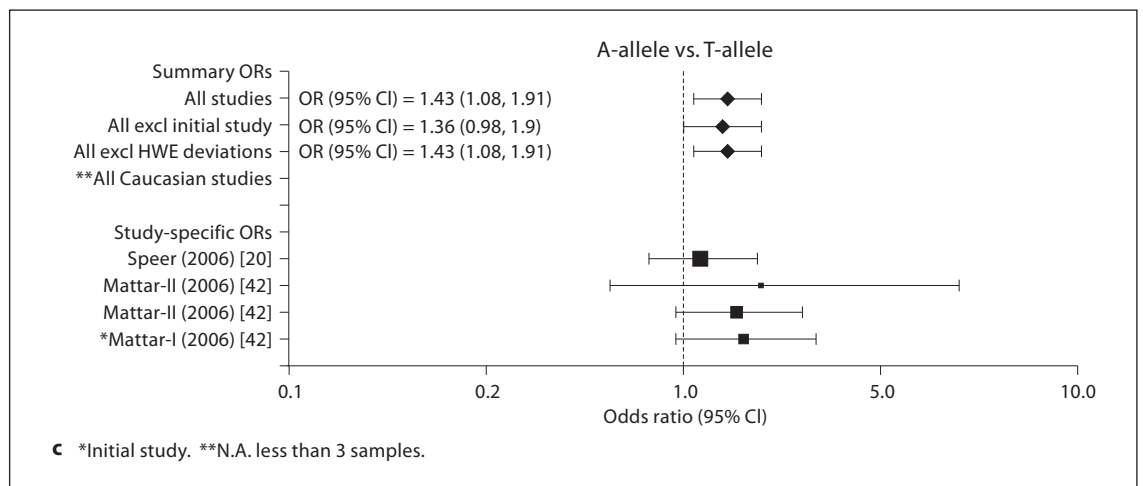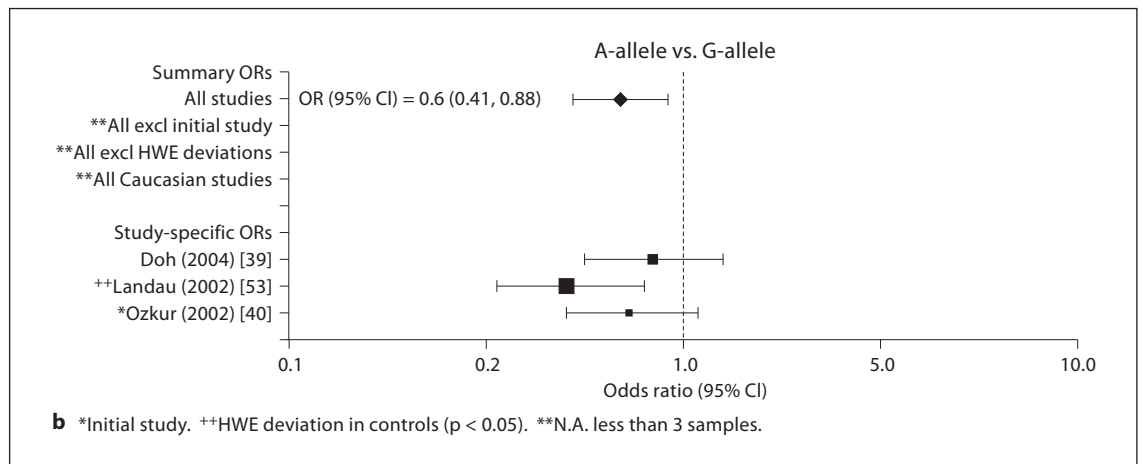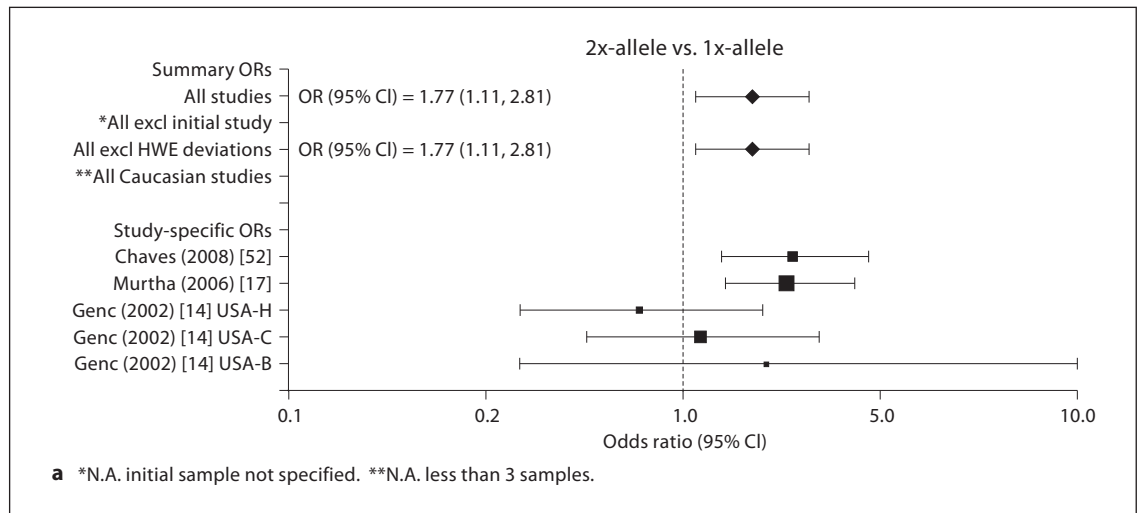
Dolan et al.

Fig. 1. Forest plots for allele-based random-effects meta-analyses on the PTBGene data as of December 1, 2008. **a** *IL1RN* variant 86 bp VNTR (maternal). **b** *ADRB2* variant rs1042713 (maternal). **c** *IFNG* variant rs2430561 (maternal).
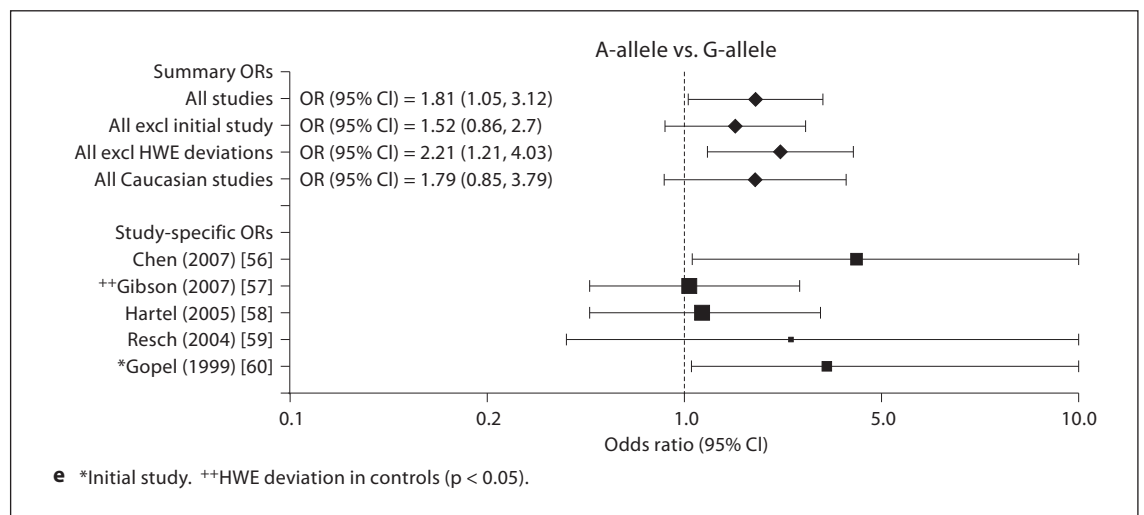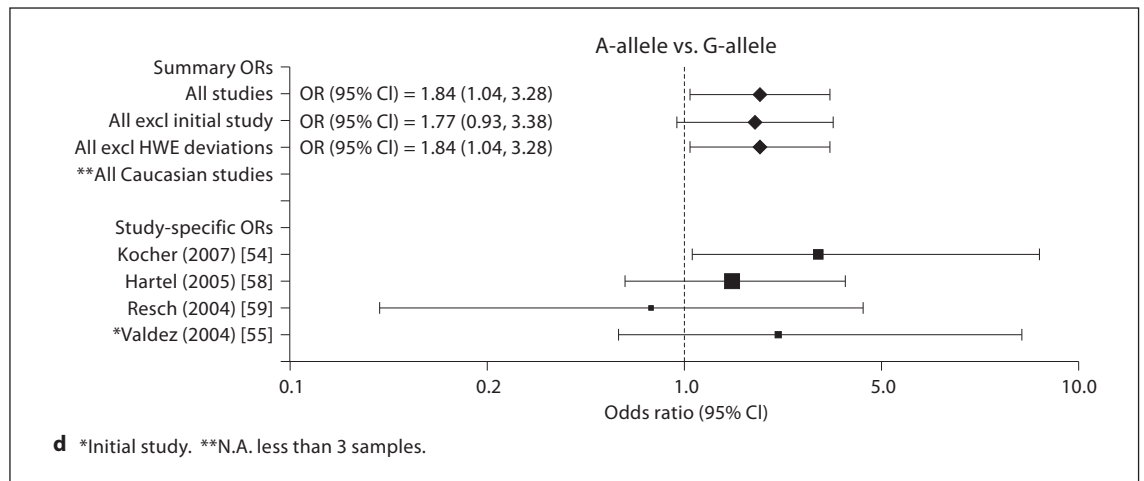
Fig. 1. Forest plots for allele-based random-effects meta-analyses on the PTBGene data as of December 1, 2008. **d** *F2* variant rs1799963 (maternal). **e** *F2* variant rs1799963 (newborn).

Table 4. Assessment of cumulative evidence [37]

Gene	Polymorphism	Amount of evidence	Replication consistency	Protection from bias	Overall
IL1RN (maternal)	86 bp VNTR	B	B	C	weak
ADRB2 (maternal)	rs1042713	B	B	C	weak
IFNG (maternal)	rs2430561	B	A	C	weak
F2 (maternal)	rs1799963	C	A	C	weak
F2 (newborn)	rs1799963	C	A	C	weak

Amount of evidence: Grade is A when the total number of minor alleles of cases and controls combined is >1,000, B when it is 100–1,000 and C when it is <100. Replication consistency: C reflects point estimates of I^2 >50%, B is 25–50%, and A is <25%. Protection from bias: Grade A means that bias, if present, may change the magnitude but not the presence of an association; grade B means that there is no evidence of bias that would invalidate an association, but important information is missing; grade C means that there is demonstrable potential or clear bias that may invalidate the mere presence of an association.

Discussion

This field synopsis of genetic associations with preterm birth reveals a paucity of research in the field, compared with more mature genetic association fields like schizophrenia [23] and Alzheimer's disease [22]. So far, no robustly replicated genetic variants contributing to this complex disease have been identified. All 5 identified nominally significant associations were graded as having weak credibility [37]. For the field to progress, the obstacles of small studies, publication and reporting biases and lack of common standards worldwide [43] need to be overcome.

Our evaluation revealed considerable problems in phenotype definition. There was large variability in the gestational cut-off used to define preterm birth. Exclusion of preterm births due to maternal or fetal indications was inconsistent. A clear phenotypic case description, according to common standards worldwide, is necessary to improve this area of research [44]. If investigators analyze associations using different cut-off definitions and report only the most impressive results, this will introduce the bias of multifarious outcomes with selective reporting [45].

This field synopsis also reveals that small studies abound in the field of preterm birth research. Most of these studies, individually and in meta-analysis, lack the power to detect the modest effect sizes observed here or in similar systematic field synopses of complex genetic diseases [22, 23]. Creation of large multicenter collaborations and consortia should be encouraged [46] with the goal of designing adequately powered studies using consistent phenotype definitions and enhancing involvement of low-resource countries. For postulated associations to date, replication efforts have been limited and even combining data in meta-analyses does not yield large enough sample sizes. Also, such small studies have limited ability to address population admixture issues as they do not have adequate power to perform ancestry-specific analyses [47–49].

Lastly, this field synopsis points out the challenges of studying an outcome such as preterm birth which is influenced by multiple genomes: the maternal, fetal and paternal genomes [38, 50]. Several triad studies are underway that may reveal genetic interactions between maternal, paternal and fetal genetic variants influencing preterm birth. However, additional methods for dealing with the task of identifying genetic associations with preterm birth, accounting for the multiple genomes, as well as environmental factors need to be developed.

We plan to keep updating this field synopsis online at PTBGene, the Preterm Birth Genetics Knowledge Base, and to re-evaluate the data on the cumulative evidence on genetic associations with preterm birth [51]. Genome-wide association studies (GWAS) are allowing high throughput hypothesis-free exploration of genetic contribution to a variety of complex outcomes. Several preterm birth GWAS expect to have results in 2010. Preterm birth research will likely benefit greatly from GWAS findings and PTBGene will incorporate data from such studies as they become available following the approach developed for AlzGene and SzGene for inclusion of large-scale genetics studies and merging them with candidate-gene data [22, 23]. Hopefully, better clinical and phenotypic characterization of study subjects, as well as larger-scale multi-center collaborations using common standards and high throughput platforms, will lead to progress in deciphering the genetic factors underlying preterm birth.

References

1 Martin JA, Hamilton BE, Sutton PD, Ventura SJ: Births: final data for 2006. Natl Vital Stat Rep 2009;57:1–101.

2 Merialdi M, Murray JC: The changing face of preterm birth. Pediatrics 2007;120:1133–1134.

3 Fawke J: Neurological outcomes following preterm birth. Semin Fetal Neonatal Med 2007;12:374–382.

4 Eriksson JG: The fetal origins hypothesis – 10 years on. BMJ 2005;330:1096–1097.

5 Lockwood CJ, Kuczynski E: Risk stratification and pathological mechanisms in preterm delivery. Paediatr Perinat Epidemiol 2001;15(suppl 2):78–89.

6 Hao K, Wang X, Niu T, Xu X, Li A, Chang W, Wang L, Li G, Laird N, Xu X: A candidate gene association study on preterm delivery: application of high-throughput genotyping technology and advanced statistical methods. Hum Mol Genet 2004;13:683–691.

7 Porter TF, Fraser AM, Hunter CY, Ward RH, Varner MW: The risk of preterm birth across generations. Obstet Gynecol 1997;90:63–67.

8 Plunkett J, Muglia LJ: Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. Ann Med 2008;40:167–195.

9 Anum EA, Springel EH, Shriver MD, Strauss JF 3rd: Genetic contributions to disparities in preterm birth. Pediatr Res 2009;65:1–9.

10 Clausson B, Lichtenstein P, Cnattingius S: Genetic influence on birthweight and gestational length determined by studies in offspring of twins. BJOG 2000;107:375–381.

11 Treloar SA, Macones GA, Mitchell LE, Martin NG: Genetic influences on premature parturition in an Australian twin sample. Twin Res 2000;3:80–82.

12 Amory JH, Adams KM, Lin MT, Hansen JA, Eschenbach DA, Hitti J: Adverse outcomes after preterm labor are associated with tumor necrosis factor-(alpha) polymorphism –863, but not –308, in mother-infant pairs. Am J Obstet Gynecol 2004;191:1362–1367.

13 Erhardt E, Stankovics J, Molnar D, Adamovich K, Melegh B: High prevalence of factor V Leiden mutation in mothers of premature neonates. Biol Neonate 2000;78:145–146.

14 Genc MR, Gerber S, Nesin M, Witkin SS: Polymorphism in the interleukin-1 gene complex and spontaneous preterm delivery. Am J Obstet Gynecol 2002;187:157–163.

15 Grisaru-Granovsky S, Tevet A, Bar-Shavit R, Salah Z, Elstein D, Samueloff A, Altarescu G: Association study of protease activated receptor 1 gene polymorphisms and adverse pregnancy outcomes: results of a pilot study in Israel. Am J Med Genet A 2007;143A:2557–2563.

16 Moore S, Ide M, Randhawa M, Walker JJ, Reid JG, Simpson NA: An investigation into the association among preterm birth, cytokine gene polymorphisms and periodontal disease. BJOG 2004;111:125–132.

17 Murtha AP, Nieves A, Hauser ER, Swamy GK, Yonish BA, Sinclair TR, Heine RP: Association of maternal IL-1 receptor antagonist intron 2 gene polymorphism and preterm birth. Am J Obstet Gynecol 2006;195:1249–1253.

18 Papazoglou D, Galazios G, Koukourakis MI, Kontomanolis EN, Maltezos E: Association of –634G/C and 936C/T polymorphisms of the vascular endothelial growth factor with spontaneous preterm delivery. Acta Obstet Gynecol Scand 2004;83:461–465.

19 Simhan HN, Krohn MA, Roberts JM, Zeevi A, Caritis SN: Interleukin-6 promoter –174 polymorphism and spontaneous preterm birth. Am J Obstet Gynecol 2003;189:915–918.

20 Speer EM, Gentile DA, Zeevi A, Pillage G, Huo D, Skoner DP: Role of single nucleotide polymorphisms of cytokine genes in spontaneous preterm delivery. Hum Immunol 2006;67:915–923.

21 Uma R, Forsyth JS, Struthers AD, Fraser CG, Godfrey V, Murphy DJ: Correlation of angiotensin converting enzyme activity and the genotypes of the I/D polymorphism in the *ACE* gene with preterm birth and birth weight. Eur J Obstet Gynecol Reprod Biol 2008;141:27–30.

22 Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE: Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet 2007; 39:17–23.

23 Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kawoura FK, Khoury MJ, Tanzi RE, Bertram L: Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. Nat Genet 2008;40:827–834.

24 Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB: Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008–2012.

25 HuGE Literature Finder. Available at http://www.hugenavigator.net/HuGENavigator/startPagePubLit.do (accessed March 12, 2010).

26 Yu W, Gwinn M, Clyne M, Yesupriya A, Khoury MJ: A navigator for human genome epidemiology. Nat Genet 2008;40:124–125.

27 Genetic Association Database (GAD). Available at http://geneticassociationdb.nih.gov/ (accessed March 12, 2010).

28 Entrez Gene. Available at http://www.ncbi.nlm.nih.gov/sites/entrez (accessed March 12, 2010).

29 Single Nucleotide Polymorphism Database (dbSNP). Available at http://www.ncbi.nlm.nih.gov/projects/SNP/ (accessed March 12, 2010).

30 Ioannidis JP, Ntzani EE, Trikalinos TA: 'Racial' differences in genetic effects for complex diseases. Nat Genet 2004;36:1312–1318.

31 Lau J, Ioannidis JP, Schmid CH: Quantitative synthesis in systematic reviews. Ann Intern Med 1997;127:820–826.

32 Ioannidis JP, Patsopoulos NA, Evangelou E: Uncertainty in heterogeneity estimates in meta-analyses. BMJ 2007;335:914–916.

33 Trikalinos TA, Salanti G, Khoury MJ, Ioannidis JP: Impact of violations and deviations in Hardy-Weinberg equilibrium on postulated gene-disease associations. Am J Epidemiol 2006;163:300–309.

34 Harbord RM, Egger M, Sterne JA: A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Stat Med 2006;25:3443–3457.

35 Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I: The case of the misleading funnel plot. BMJ 2006;333:597–600.

36 Ioannidis JP, Trikalinos TA: An exploratory test for an excess of significant findings. Clin Trials 2007;4:245–253.

37 Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, Balding DJ, Chokkalingam A, Dolan SM, Flanders WD, Higgins JP, McCarthy MI, McDermott DH, Page GP, Rebbeck TR, Seminara D, Khoury MJ: Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol 2008;37:120–132.

38 Lie RT, Wilcox AJ, Skjaerven R: Maternal and paternal influences on length of pregnancy. Obstet Gynecol 2006;107:880–885.

39 Doh K, Sziller I, Vardhana S, Kovacs E, Papp Z, Witkin SS: Beta2-adrenergic receptor gene polymorphisms and pregnancy outcome. J Perinat Med 2004;32:413–417.

40 Ozkur M, Dogulu F, Ozkur A, Gokmen B, Inaloz SS, Aynacioglu AS: Association of the Gln27Glu polymorphism of the beta-2-adrenergic receptor with preterm labor. Int J Gynaecol Obstet 2002;77:209–215.

41 Schroder K, Hertzog PJ, Ravasi T, Huma DA: Interferon-gamma: an overview of signals, mechanisms and functions. J Leukoc Biol 2004;75:163–189.

42 Mattar R, De Souza E, Daher S: Preterm delivery and cytokine gene polymorphisms. J Reprod Med 2006;51:317–320.

43 Ioannidis JP, Bernstein J, Boffetta P, Danesh J, Dolan S, Hartge P, Hunter D, Inskip P, Jarvelin MR, Little J, Maraganore DM, Bishop JA, O'Brien TR, Petersen G, Riboli E, Seminara D, Taioli E, Uitterlinden AG, Vineis P, Winn DM, Salanti G, Higgins JP, Khoury MJ: A network of investigator networks in human genome epidemiology. Am J Epidemiol 2005;162:302–304.

44 Pennell CE, Jacobsson B, Williams SM, Buus RM, Muglia LJ, Dolan SM, Morken NH, Ozcelik H, Lye SJ, Relton C: Genetic epidemiologic studies of preterm birth: guidelines for research. Am J Obstet Gynecol 2007;196: 107–118.

45 Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JP: An empirical evaluation of multifarious outcomes in pharmacogenetics: beta-2 adrenoceptor gene polymorphisms in asthma treatment. Pharmacogenet Genomics 2006;16:705–711.

46 Seminara D, Khoury MJ, O'Brien TR, Manolio T, Gwinn ML, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, Bray MS, Brenchley PE, Buffler PA, Casas JP, Chokkalingam AP, Danesh J, Davey-Smith G, Dolan S, Duncan R, Gruis NA, Hashibe M, Hunter D, Jarvelin MR, Malmer B, Maraganore DM, Newton-Bishop JA, Riboli E, Salanti G, Taioli E, Timpson N, Uitterlinden AG, Vineis P, Wareham N, Winn DM, Zimmern R, Ioannidis JP; Human Genome Epidemiology Network; Network of Investigator Networks: The emergence of networks in human genome epidemiology: challenges and opportunities. Epidemiology 2007;18:1–8.

47 Menon R, Fortunato SJ, Thorsen P, Williams S: Genetic associations in preterm birth: a primer of marker selection, study design, and data analysis. J Soc Gynecol Investig 2006;13:531–541.

48 Menon R, Velez DR, Thorsen P, Vogel I, Jacobsson B, Williams SM, Fortunato SJ: Ethnic differences in key candidate genes for spontaneous preterm birth: TNF-alpha and its receptors. Hum Hered 2006;62:107–118.

49 Velez DR, Menon R, Thorsen P, Jiang L, Simhan H, Morgan N, Fortunato SJ, Williams SM: Ethnic differences in interleukin 6 (IL-6) and *IL6* receptor genes in spontaneous preterm birth and effects on amniotic fluid protein levels. Ann Hum Genet 2007;71: 586–600.

50 Wilcox AJ, Skjaerven R, Lie RT: Familial patterns of preterm delivery: maternal and fetal contributions. Am J Epidemiol 2008;167: 474–479.

51 Ioannidis JP, Gwinn M, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, Bray MS, Brenchley PE, Buffler PA, Casas JP, Chokkalingam A, Danesh J, Smith GD, Dolan S, Duncan R, Gruis NA, Hartge P, Hashibe M, Hunter DJ, Jarvelin MR, Malmer B, Maraganore DM, Newton-Bishop JA, O'Brien TR, Petersen G, Riboli E, Salanti G, Seminara D, Smeeth L, Taioli E, Timpson N, Uitterlinden AG, Vineis P, Wareham N, Winn DM, Zimmern R, Khoury MJ: A road map for efficient and reliable human genome epidemiology. Nat Genet 2006;38:3–5.

52 Chaves JH, Babayan A, Bezerra Cde M, Linhares IM, Witkin SS: Maternal and neonatal interleukin-1 receptor antagonist genotype and pregnancy outcome in a population with a high rate of pre-term birth. Am J Reprod Immunol 2008;60:312–317.

53 Landau R, Xie HG, Dishy V, Stein CM, Wood AJJ, Moore JH, Emala CW, Smiley RM: Beta2-adrenergic receptor genotype and preterm delivery. Am J Obstet Gynecol 2002; 187:1294–1298.

54 Kocher O, Cirovic C, Malynn E, Rowland CM, Bare LA, Young BA, Henslee JG, Laffler TG, Huff JB, Kruskall MS, Wong G, Bauer KA: Obstetric complications in patients with hereditary thrombophilia identified using the LCx microparticle enzyme immunoassay: a controlled study of 5,000 patients. Am J Clin Pathol 2007;127:68–75.

55 Valdez LL, Quintero A, Garcia E, Olivares N, Celis A, Rivas F Jr, Rivas F: Thrombophilic polymorphisms in preterm delivery. Blood Cells Mol Dis 2004;33:51–56.

56 Chen BH, Carmichael SL, Shaw GM, Iovannisci DM, Lammer EJ: Association between 49 infant gene polymorphisms and preterm delivery. Am J Med Genet A 2007;143:1990–1996.

57 Gibson CS, MacLennan AH, Dekker GA, Goldwater PN, Dambrosia JM, Munroe DJ, Tsang S, Stewart C, Nelson KB: Genetic polymorphisms and spontaneous preterm birth. Obstet Gynecol 2007;109:384–391.

58 Hartel C, von Otte S, Koch J, Ahrens P, Kattner E, Segerer H, Moller J, Diedrich K, Gopel W: Polymorphisms of haemostasis genes as risk factors for preterm delivery. Thromb Haemost 2005;94:88–92.

59 Resch B, Gallistl S, Kutschera J, Mannhalter C, Muntean W, Mueller WD: Thrombophilic polymorphisms – Factor V Leiden, prothrombin G20210A, and methylenetetrahydrofolate reductase C677T mutations – and preterm birth. Wien Klin Wochenschr 2004;116:622–626.

60 Gopel W, Kim D, Gortner L: Prothrombotic mutations as a risk factor for preterm birth. Lancet 1999;353:1411–1412.