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

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

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

Key Words
HuGE review · Meta-analysis · Prematurity · Preterm birth · Systematic review

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.
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

<table>
<thead>
<tr>
<th>Databases: EMBASE + MEDLINE</th>
<th>Limited to human</th>
<th>Years: January 1, 1990 to December 1, 2008</th>
</tr>
</thead>
</table>

- **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 birth’ 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 ‘premature birth’ or ‘pre-mature bab’ or ‘preterm bab’ or ‘premature 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**
  - ‘premature’ or ‘pre-mature’ 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 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 I² 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-

Table 2. Systematic review

<table>
<thead>
<tr>
<th>Overall data set</th>
<th>Maternal data</th>
<th>Newborn data</th>
<th>Paternal data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of articles</td>
<td>61</td>
<td>46</td>
<td>31</td>
</tr>
<tr>
<td>Number of genes</td>
<td>84</td>
<td>48</td>
<td>68</td>
</tr>
<tr>
<td>Number of polymorphisms</td>
<td>189</td>
<td>127</td>
<td>145</td>
</tr>
</tbody>
</table>

6,772 abstracts were identified and screened. 139 full-text articles were retrieved. 61 articles had data extracted.
erogeneity and values of 25–50% express moderate between-study heterogeneity. With a small number of studies, I² 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 (ILIRN) 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 \( \sim 1.4 \) to \( \sim 1.8 \), \( p \) 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-
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 9H252-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. I² 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).

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 | Heterogeneity | 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 | I² = 36 | p = 0.18 | 211 | 5 |
| ADRB2 (maternal) | rs1042713 A vs. G | 0.60 (0.41–0.88) | p = 0.01 | I² = 28 | p = 0.25 | 516 | 3 |
| IFNG (maternal) | rs2430561 A vs. T | 1.43 (1.08–1.91) | p = 0.01 | I² = 0 | p = 0.71 | 508 | 4 |
| F2 (maternal) | rs1799963 A vs. G | 1.84 (1.04–3.28) | p = 0.04 | I² = 0 | p = 0.55 | 51 | 4 |
| F2 (newborn) | rs1799963 A vs. G | 1.81 (1.05–3.12) | p = 0.03 | I² = 15 | p = 0.32 | 75 | 5 |
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).
Summary ORs

All studies
OR (95% CI) = 1.84 (1.04, 3.28)

All excl initial study
OR (95% CI) = 1.77 (0.93, 3.38)

All excl HWE deviations
OR (95% CI) = 1.84 (1.04, 3.28)

**All Caucasian studies

Study-specific ORs
Kocher (2007) [54]
Hartel (2005) [58]
Resch (2004) [59]
*Gopel (1999) [60]

0.1
0.2
1.0
5.0
10.0

Odds ratio (95% CI)

Fig. 1. Forest plots for allele-based random-effects meta-analyses on the PTBGene data as of December 1, 2008.

F2 variant rs1799963 (maternal).
F2 variant rs1799963 (newborn).

Table 4. Assessment of cumulative evidence [37]

<table>
<thead>
<tr>
<th>Gene</th>
<th>Polymorphism</th>
<th>Amount of evidence</th>
<th>Replication consistency</th>
<th>Protection from bias</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL1RN (maternal)</td>
<td>86 bp VNTR</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>weak</td>
</tr>
<tr>
<td>ADRB2 (maternal)</td>
<td>rs1042713</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>weak</td>
</tr>
<tr>
<td>IFNG (maternal)</td>
<td>rs2430561</td>
<td>B</td>
<td>A</td>
<td>C</td>
<td>weak</td>
</tr>
<tr>
<td>F2 (maternal)</td>
<td>rs1799963</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>weak</td>
</tr>
<tr>
<td>F2 (newborn)</td>
<td>rs1799963</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>weak</td>
</tr>
</tbody>
</table>

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 I² >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.
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.

Acknowledgements

The authors thank I. Merkatz, M. Katz, C. Hobel and the HuGEnet and PREBIC memberships, J. Kinoshita and the Alzheimer Research Forum, the Genetics and Aging Research Unit at Massachusetts General Hospital, M. Gwinn, W. Yu and the Office of Public Health Genomics at CDC, and V. Niranjan at Albert Einstein College of Medicine, for technical assistance.

Funding for this project has been provided by the World Health Organization, Preterm Birth International Collaborative, Albert Einstein College of Medicine, March of Dimes, and the Cure Alzheimer’s Fund.

This study was designed by S.M.D., M.J.K., L.B., and J.P.I. Literature searches were performed by T.A. Articles were reviewed and data extracted by C.A. and PREBIC members (S.M.D., M.V.H., M.M., A.P.B., J.N., B.L., D.R.V., R.M.). Online curation of data was performed by C.A. and S.M.D. Analysis was performed by L.B., J.P.I. and S.M.D. The manuscript was written by S.M.D. and J.P.I. with contributions from M.J.K., M.V.H., M.M., D.R.V., L.B. and R.M. The website was developed by L.B., X.Z., R.A.D. and S.M.D. The design and layout of the online version of PTBGene was modeled after the layout developed for AlzGene and related databases by Massachusetts General Hospital (L.B.) and the Alzheimer Research Forum.
References


