

Genes and Environment in Common Neonatal Lung Disease

Mikko Hallman^a Riitta Marttila^b Riccardo Pertile^c Marja Ojaniemi^a
Ritva Haataja^a

^aDepartment of Pediatrics, Biocenter Oulu, University of Oulu, Oulu, ^bCentral Hospital of Southern Ostrobothnia, Seinäjoki, Finland; ^cUniversity of Padua, Padua, Italy

Key Words

Respiratory distress syndrome · Bronchopulmonary dysplasia · Preterm birth · Prematurity · Genetics · Genome-wide screening · Biobank

Abstract

Respiratory distress syndrome (RDS) and bronchopulmonary dysplasia (BPD) are common, serious lung diseases in preterm infants. Polymorphism of the genes involved in basic lung function and alveolar stability, lung differentiation and pulmonary host defense may influence the risk. Natural selection has refined the genes responsible for cardiopulmonary adaptation and resistance against pneumonia in term and near-term infants. Before the era of antibiotics, however, virtually all very preterm infants died of asphyxia, respiratory failure or infections. Today, the degree of prematurity plays a dominant role in susceptibility to serious lung disease. In addition, genetic polymorphism and constitution modulate the risk of RDS and BPD that have different, partly overlapping predisposition. According to twin studies, the genetic impact on the risk of RDS and BPD among preterm and very preterm infants is 35–65%. Individual disease genes generally have low penetrance. Large-scale genetic studies are required as part of neonatal and perinatal research in order to learn about the risk factors and to investigate pharmacogenetics. The aim in the future is to individualize therapies.

Copyright © 2007 S. Karger AG, Basel

Introduction

Perinatal respiratory adaptation is an effective sequence of complex events. However, respiratory failure is the most common cause of death in early infancy. Successful respiratory adaptation in near-term infants has been an evolutionary advantage. Previously, however, the population of very prematurely born infants had additional lethal diseases in the event of them not dying of asphyxia or respiratory distress syndrome (RDS) shortly after birth. As a result of new prenatal and neonatal treatments, the prognosis of these infants has improved dramatically. Because the limit of viability has been lowered by 8–10 weeks within 50 years, prematurity continues to be the prominent basic cause of perinatal and infant death and the major cause of chronic disability, particularly due to bronchopulmonary dysplasia (BPD) and neurosensory diseases.

RDS presents as a transient deficiency of alveolar surfactant that is influenced by polymorphisms of the genes playing a critical role in homeostasis of type 2 alveolar cells. Apart from surfactant deficiency, a number of interactive constitutional, environmental and genetic factors disturb neonatal respiratory adaptation and delay the recovery from RDS, and may thus influence the genetic predisposition. The susceptibility to BPD is likely to be affected by polymorphisms of genes influencing a number of critical pulmonary functions ranging from

the maintenance of gas exchange to host defense, lung growth and differentiation.

The susceptibility of very preterm infants to RDS or BPD has not been directly under the selection pressure of evolution, as all these infants used to die early. It is therefore not very surprising to observe that the genotypes increasing the risk have generally a high frequency. In the present brief review, we discuss the genetic susceptibility to common lung diseases after premature birth.

Diagnosis of Lung Disease in Small Preterm Infants

The definition of lung diseases in very preterm (VLGA; <32 weeks of gestation) infants is not always accurate. However, it is important to systematically define these diseases and to try to improve diagnostic methods.

BPD is prevalent among VLGA infants, especially among those born extremely preterm (ELGA; <28 weeks of gestation). BPD in VLGA infants is defined as a generalized lung disease requiring supplemental oxygen at 28 days of age (CLD or 'mild' BPD) and at 36 weeks of post-menstrual age (moderate to severe BPD) [1, 2]. A longer follow-up may give a more accurate definition, as some infants with BPD have low morbidity later in infancy. Others many suffer from recurrent hospitalizations due to lung disease, which contributes to delay in growth and neurodevelopment.

The diagnosis of RDS has become problematic due to the various treatments available at or after birth. Surfactant therapy in the delivery room is considered an exclusion criterion for genetic studies of RDS, since this treatment may conceal the characteristic symptoms and diagnostic changes in chest x-rays. On the other hand, antenatal glucocorticoid treatment, which decreases the risk of RDS, is a confounder rather than a contraindication for genetic studies. The major haplotype of *SP-A1* 6A2 is a risk factor among VLGA infants. This haplotype of *SP-A* is very prominent among those VLGA infants who develop RDS despite antenatal glucocorticoid treatment, and is underrepresented among those VLGA infants with neither antenatal corticosteroid treatment nor RDS [3].

Other Lung Diseases

The susceptibility to other neonatal pulmonary diseases, including infections or meconium aspiration syndrome, is also likely to have a genetic component. Ge-

netic susceptibility to severe respiratory syncytial virus infection during epidemics has been studied [4, 5]. In mice, the susceptibility to group B β -hemolytic streptococcal (GBS) pneumonia appears to have a genetic background [6] and genetic predisposition to other neonatal pulmonary infections is likely [7]. However, very few studies have been published so far.

Differentiation between Genetic and Environmental Factors

The close association between genetic and environmental factors complicates the evaluation of any genetic impact on a risk of disease. Another factor affecting the evaluation is the extreme phenotype. For instance, the generalized multiorgan immaturity in ELGA infants born before 23–25 weeks' gestation is due to their extreme vulnerability and lack of adequate host defense. RDS is virtually always diagnosed in these cases. In addition, BPD and diseases involving the central nervous system, the cardiovascular system and the gastrointestinal tract are prevalent. Inclusion of these cases may decrease the penetrance of the disease genes or result in an evaluation of the genetic risk of extreme prematurity rather than the associated diseases. Genetic association with RDS or BPD may become detectable first when the population is limited on the basis of the degree of prematurity or of other environmental stress factors.

RDS and BPD are separate, partly overlapping clinical entities. However, the genetic and environmental risk factors and their interactions are not likely to be homogenous. The phenotype of RDS in the population of near-term infants is distinguishable from that among VLGA infants. The alleles, genotypes and haplotypes associating with RDS in near-term infants are mostly different from the alleles associating with RDS in VLGA infants. The disease genotypes in near-term infants tend to have a low frequency and rather high penetrance [8], whereas the disease genotypes in the RDS of VLGA infants tend to have a high frequency and low penetrance [3, 9, 10]. In contrast, certain rare mutations with very high penetrance are associated with fatal respiratory failure in mostly term-born infants [11].

The well-known environmental (e.g. gestational age) and constitutional (e.g. gender) factors influencing the risk of the lung disease are usually analyzed as independent variables. An association between the genotype, the risk factor (or protective factor) and the phenotype may, at best, give a clue to the function of the disease genotype.

Family Studies. Studies on sibs or other relatives help to link the common hereditary factors with the risk of disease, providing a numerical estimate of heritability. Genome-wide linkage analysis of large families with a high risk of the disease could optimally define the linkage between the disease and a DNA locus. However, it would be difficult to recruit a sufficient study population, since these diseases are predominant only in preterm infants. Although the recurrence rate of preterm birth increases 3-fold after a single premature birth and 6-fold after two premature births, the variation in the degree of prematurity complicates the assessment of the risk. Inclusion of the families with recurrent prematurity will limit the population to cases with a genetic predisposition to preterm birth. In twin pregnancies, on the other hand, the environmental and constitutional differences appear to be small, and the difference in consanguinity between dizygotic (DZ) and monozygotic (MZ) twin pairs (50% among DZ twin pairs, 100% among MZ) consolidates the evaluation of heritability [12]. Thus, the concordance difference between MZ and DZ twins provides a direct estimate of the heritability of the disease. Twin studies have yielded evidence indicating that the genetic risks of BPD [13] and RDS [14] are considerable, with a wide range in estimates.

There are caveats in twin studies, however. A difference in the mean gestational age between MZ and DZ twin pairs and the weight or gender difference within twin pairs perturbs the concordance difference estimates. Moreover, the calculated heritability figure may not be specific, since unidentified environmental factors disturb the concordance difference. For instance, leukemia cells may be transmitted non-genetically to a MZ twin pair via placental anastomoses [15]. Likewise, the multiple pregnancy setting perturbs the concordance of RDS, as the presenting twin has a lower risk of RDS than the non-presenting twin. The acceleration of lung maturation of the presenting twin is under genetic control that involves the *SP-B* Ile131Thr polymorphism [16]. The low RDS risk is confined to the presenting fetuses, who were carriers of the common *SP-B* Ile131 allele, whereas the RDS risk of the non-presenting fetus is not influenced by the *SP-B* Ile131Thr genotypes.

Studies Identifying the Genetic Locus

Studies on genetic susceptibility generally require a large population sample because of the abundance of dependent variables and the generally low penetrance of genetic variants. The power estimate is mandatory, and the

limits of power need to be realized. Prospectively designed case-control studies are preferred whenever the cases and controls show nearly equal population representation, and the major dependent variables (at least the length of gestation) can be equated. On the other hand, population cohorts with minimal selection bias make it possible to evaluate the overall impact of genetic factors. In post-hoc analysis, the independent variables need to be carefully considered. Penetrance, zygosity and genotype frequency influence the risk assessment.

The genetic background of the population requires consideration. The difference in frequencies of individual genotypes as compared between the different populations may influence the genetic risk in a more complex way than anticipated on the basis of the different frequencies (effect of genetic constitution). Therefore, ethnic/racial variation needs to be considered independently. Genetically homogenous populations, particularly when there is evidence on a recent confounder effect [17], are suitable for the identification of a genetic risk. Generally, a smaller sample of a homogenous than a heterogenous population is required for the study.

To ascertain the heredity of a trait, inclusion of the family triad (mother, father and child) as an addition to the case-control setting is preferred. This allows examination of the linkage between the gene and the disease using the transmission disequilibrium test (TDT). In the TDT the aim is to evaluate whether the transmission of the putative disease genotype to individuals with the disease is significantly favored. It is also possible to study the transmission of the protective genotype to healthy 'supernormal' individuals (i.e. healthy despite environmental stress) [18].

Demonstration of a genetic association with the disease requires both case-control studies and those involving studies of genetic linkage (such as TDT). However, positive association needs to be confirmed in other independent studies. Discrepancies in the results are not uncommon, and meta-analyses are required to judge the overall significance of the genetic association.

Candidate Genes

Studies on the pathogenesis of a disease help to define the candidate genes. The evidence on lethal neonatal respiratory failure caused by the lack-of-function mutation in a gene serves as an indication for studies on the gene polymorphism that may increase the risk of multifactorial lung disease (RDS or BPD). Such evidence has been

Table 1. Candidate genes that have a proposed association with susceptibility to RDS and BPD

| Disease/gene | Risk (protective) allele or haplotype | References |
|--------------|---------------------------------------|-------------------|
| RDS | | |
| <i>SP-A</i> | SP-A1 6A2 (6A3) haplotype | 3, 9, 10, 20 |
| <i>SP-B</i> | 131Thr | 9, 10, 16, 21, 22 |
| <i>SP-C</i> | 186Asn | 23, 24 |
| <i>GPRA</i> | H4/5 (H1) haplotype | 8 |
| <i>ABCA3</i> | rs13332514 in exon 10 | 25 |
| <i>IFN-γ</i> | 874T | 26 |
| BPD | | |
| <i>SP-B</i> | i4del, or various alleles | 24, 27 |
| <i>ABCA3</i> | rs13332514 in exon 10 | 25 |
| <i>TNF-α</i> | (-308A), or no association | 28, 29 |
| <i>IFN-γ</i> | (874T) | 26 |
| <i>ACE</i> | Deletion allele | 30 |

obtained for at least *SP-B* and *ABCA3*, and several other genes critical for lung function and preferentially expressed in the lung may become interesting targets for research on multifactorial lung diseases [19]. Haplotype analysis yields information about DNA stretches with an exceptionally predictable set of sequences that are in close linkage disequilibrium. Common alleles, haplotypes and genotypes of representative genes are often studied. On the other hand, rare alleles and genotypes may occasionally be significant risk factors even of common multifactorial lung diseases, provided that the detrimental effect of the allele has a high penetrance.

Table 1 lists many of the genes and alleles proposed as candidate genes for RDS and BPD. All studies are preliminary or controversial, and all associations need to be confirmed. In most studies the association is limited to a phenotype that is restricted by the length of gestation at birth.

Genomic Studies

The microsatellites that contain variable repeat sequences of 2–4 base pairs have been commonly used in linkage analysis. Other techniques are increasingly applied. The microchip techniques currently allow the analysis of single nucleotide polymorphisms (SNP) in the range of 1 million. In molecular biology these genomic and proteomic techniques are already commonly applied. In common diseases of adults (diabetes, cardiovascular

diseases) genomic approaches are increasingly used and will likely be applied in neonatal medicine as well. The current price of a single SNP (about 10 cents) is perhaps misleading, since besides the association study, a remarkable amount of additional research is required to define a single disease genotype and its function. Selective genomic analysis involving genes with similar functional category (e.g. innate immunity) are considered as well.

Comment

In view of the aim to increase quality-adjusted life years of sick newborn infants, research on genetics of common neonatal diseases is justified. However, such research projects require prioritization, dedicated networks, multidisciplinary teams and sufficient resources. A biobank linked with a database would, in theory, provide accurate genomic, constitutional and hospital data. However, the ethical issues, legislation and acceptance within the community need to be overcome. Computational research and large projects involving translational and clinical research would be required before approaching potential applications. The aim to understand and define the disease genes influencing susceptibility to neonatal lung diseases should not supersede the continuous research and development in other aspects of neonatal medicine. Despite possible delays and problems, genetic studies are likely to expand and diversify. We predict that eventually they will involve trials with pharmacogenetic incentives to develop a new generation of individualized therapies.

References

- 1 Jobe AH, Bancalari E: Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001; 163:1723–1729.
- 2 Walsh MC, Yao Q, Gettner P, Hale E, Collins M, Hensman A, Everette R, Peters N, Miller N, Muran G, Auten K, Newman N, Rowan G, Grisby C, Arnell K, Miller L, Ball B, McDavid G, National Institute of Child Health and Human Development Neonatal Research Network: Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics* 2004;114:1305–1311.
- 3 Ramet M, Haataja R, Marttila R, Floros J, Hallman M: Association between the surfactant protein A gene locus and respiratory-distress syndrome in the Finnish population. *Am J Hum Genet* 2000;66:1569–1579.

- 4 Lofgren J, Ramet M, Renko M, Marttila R, Hallman M: Association between surfactant protein A gene locus and severe respiratory syncytial virus infection in infants. *J Infect Dis* 2002;185:283–289.
- 5 Lahti M, Lofgren J, Marttila R, Renko M, Klaavuniemi T, Haataja R, Ramet M, Hallman M: Surfactant protein D gene polymorphism associated with severe respiratory syncytial virus infection. *Pediatr Res* 2002; 51:696–699.
- 6 Mancuso G, Midiri A, Beninati C, Biondo C, Galbo R, Akira S, Henneke P, Golenbock D, Teti G: Dual role of TLR2 and myeloid differentiation factor 88 in a mouse model of invasive group B streptococcal disease. *J Immunol* 2004;172:6324–6329.
- 7 Hartel C, Konig I, Koster S, Kattner E, Kuhls E, Kuster H, Moller J, Muller D, Kribs A, Segerer H, Wieg C, Herting E, Gopel W: Genetic polymorphisms of hemostasis genes and primary outcome of very low birth weight infants. *Pediatrics* 2006;118:683–689.
- 8 Pulkkinen V, Haataja R, Hannelius U, Helve O, Pitkanen OM, Karikoski R, Rehn M, Marttila R, Lindgren CM, Hastbacka J, Andersson S, Kere J, Hallman M, Laitinen T: G protein-coupled receptor for asthma susceptibility associates with respiratory distress syndrome. *Ann Med* 2006;38:357–366.
- 9 Haataja R, Ramet M, Marttila R, Hallman M: Surfactant proteins A and B as interactive genetic determinants of neonatal respiratory distress syndrome. *Hum Mol Genet* 2000;9: 2751–2760.
- 10 Marttila R, Haataja R, Guttentag S, Hallman M: Surfactant protein A and B genetic variants in respiratory distress syndrome in singletons and twins. *Am J Respir Crit Care Med* 2003;168:1216–1222.
- 11 Noguee LM: Genetic mechanisms of surfactant deficiency. *Biol Neonate* 2004;85:314–318.
- 12 Teikari JM, Kaprio J, Koskenvuo MK, Vannas A: Heritability estimate for refractive errors – a population-based sample of adult twins. *Genet Epidemiol* 1988;5:171–181.
- 13 Bhandari V, Bizzarro MJ, Shetty A, Zhong X, Page GP, Zhang H, Ment LR, Gruen JR, Neonatal Genetics Study Group: Familial and genetic susceptibility to major neonatal morbidities in preterm twins. *Pediatrics* 2006; 117:1901–1906.
- 14 Hallman M, Haataja R: Surfactant protein polymorphisms and neonatal lung disease. *Semin Perinatol* 2006;30:350–361.
- 15 Greaves MF, Maia AT, Wiemels JL, Ford AM: Leukemia in twins: lessons in natural history. *Blood* 2003;102:2321–2333.
- 16 Marttila R, Haataja R, Ramet M, Lofgren J, Hallman M: Surfactant protein B polymorphism and respiratory distress syndrome in premature twins. *Hum Genet* 2003;112:18–23.
- 17 Varilo T, Paunio T, Parker A, Perola M, Meyer J, Terwilliger JD, Peltonen L: The interval of linkage disequilibrium detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Hum Mol Genet* 2003;12:51–59.
- 18 Haataja R, Marttila R, Uimari P, Lofgren J, Ramet M, Hallman M: Respiratory distress syndrome: evaluation of genetic susceptibility and protection by transmission disequilibrium test. *Hum Genet* 2001;109:351–355.
- 19 Maeda Y, Dave V, Whitsett JA: Transcriptional control of lung morphogenesis. *Physiol Rev* 2007;87:219–244.
- 20 Kala P, Ten Have T, Nielsen H, Dunn M, Floros J: Association of pulmonary surfactant protein A (SP-A) gene and respiratory distress syndrome: interaction with SP-B. *Pediatr Res* 1998;43:169–177.
- 21 Floros J, Veletza SV, Kotikalapudi P, Krizkova L, Karinch AM, Friedman C, Buchter S, Marks K: Dinucleotide repeats in the human surfactant protein B gene and respiratory distress syndrome. *Biochem J* 1995;305:583–590.
- 22 Floros J, Thomas NJ, Liu W, Papagaroufalos C, Xanthou M, Pereira S, Fan R, Guo X, Diangelo S, Pavlovic J: Family-based association tests suggest linkage between surfactant protein B (SP-B) (and flanking region) and respiratory distress syndrome (RDS): SP-B haplotypes and alleles from SP-B-linked loci are risk factors for RDS. *Pediatr Res* 2006;59: 616–621.
- 23 Lahti M, Marttila R, Hallman M: Surfactant protein C gene variation in the Finnish population – association with perinatal respiratory disease. *Eur J Hum Genet* 2004;12:312–320.
- 24 Rova M, Haataja R, Marttila R, Ollikainen V, Tammela O, Hallman M: Data mining and multiparameter analysis of lung surfactant protein genes in bronchopulmonary dysplasia. *Hum Mol Genet* 2004;13:1095–1104.
- 25 Karjalainen M, Haataja R, Hallman M: Tagging SNP selection and haplotype analysis of ABCA3 gene polymorphisms in relation to respiratory distress syndrome in preterm infants. *Proc Am Thoracic Soc* 2007, in press.
- 26 Bokodi G, Derzbach L, Banyasz I, Tulassay T, Vasarhelyi B: Association of interferon- γ T+874A and interleukin-12 p40 promoter CTCTAA/GC polymorphism with the need for respiratory support and perinatal complications in low birthweight neonates. *Arch Dis Child Fetal Neonatal Ed* 2007;92:F25–F29.
- 27 Makri V, Hospes B, Stoll-Becker S, Borkhardt A, Gortner L: Polymorphisms of surfactant protein B encoding gene: modifiers of the course of neonatal respiratory distress syndrome? *Eur J Pediatr* 2002;161:604–608.
- 28 Kazzi SN, Kim UO, Quasney MW, Buhimschi I: Polymorphism of tumor necrosis factor- α and risk and severity of bronchopulmonary dysplasia among very low birth weight infants. *Pediatrics* 2004;114:e243–e248.
- 29 Adcock K, Hedberg C, Loggins J, Kruger TE, Baier RJ: The TNF- α -308, MCP-1 -2518 and TGF- β 1 +915 polymorphisms are not associated with the development of chronic lung disease in very low birth weight infants. *Genes Immun* 2003;4:420–426.
- 30 Kazzi SN, Quasney MW: Deletion allele of angiotensin-converting enzyme is associated with increased risk and severity of bronchopulmonary dysplasia. *J Pediatr* 2005;147: 818–822.