Gene Expression Profiling in Preterm Infants: New Aspects of Bronchopulmonary Dysplasia Development

Jacek J. Pietrzyk¹ , Przemko Kwinta1*, Embjorg J. Wollen² , Mirosław Bik-Multanowski³ , Anna Madetko-Talowska³ , Clara-Cecilie Günther⁴ , Mateusz Jagła¹ , Tomasz Tomasik¹ , Ola D. Saugstad²

1 Department of Pediatrics, Jagiellonian University, Krakow, Poland, **2** Department of Pediatric Research, University of Oslo, Oslo, Norway, **3** Department of Medical Genetics, Jagiellonian University, Krakow, Poland, **4** Norwegian Computing Center, Oslo, Norway

Abstract

Rationale: Bronchopulmonary dysplasia is one of the most serious complications observed in premature infants. Thanks to microarray technique, expression of nearly all human genes can be reliably evaluated.

Objective: To compare whole genome expression in the first month of life in groups of infants with and without bronchopulmonary dysplasia.

Methods: 111 newborns were included in the study. The mean birth weight was 1029g (SD:290), and the mean gestational age was 27.8 weeks (SD:2.5). Blood samples were drawn from the study participants on the 5th, 14th and 28th day of life. The mRNA samples were evaluated for gene expression with the use of GeneChip® Human Gene 1.0 ST microarrays. The infants were divided into two groups: bronchopulmonary dysplasia (n=68) and control (n=43).

Results: Overall 2086 genes were differentially expressed on the day 5, only 324 on the day 14 and 3498 on the day 28. Based on pathway enrichment analysis we found that the cell cycle pathway was up-regulated in the bronchopulmonary dysplasia group. The activation of this pathway does not seem to be related with the maturity of the infant. Four pathways related to inflammatory response were continuously on the $5th$, 14th and 28th day of life down-regulated in the bronchopulmonary dysplasia group. However, the expression of genes depended on both factors: immaturity and disease severity. The most significantly down-regulated pathway was the T cell receptor signaling pathway.

Conclusion: The results of the whole genome expression study revealed alteration of the expression of nearly 10% of the genome in bronchopulmonary dysplasia patients.

Citation: Pietrzyk JJ, Kwinta P, Wollen EJ, Bik-Multanowski M, Madetko-Talowska A, et al. (2013) Gene Expression Profiling in Preterm Infants: New Aspects of Bronchopulmonary Dysplasia Development . PLoS ONE 8(10): e78585. doi:10.1371/journal.pone.0078585

Editor: Susanne Krauss-Etschmann, Ludwig-Maximilians-University Munich, Germany

Received December 7, 2011; **Accepted** September 23, 2013; **Published** October 23, 2013

Copyright: © 2013 Pietrzyk et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by unrestricted grants of Financial Mechanism of European Economic Area (PL0226) and Polish Ministry of Science (E023/P01/2008/02/85). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: kwintap@mp.pl

Introduction

Bronchopulmonary dysplasia (BPD) is a chronic lung disease associated with premature birth and characterized by early lung injury [[1](#page-9-0)]. The current consensus is that BPD is a complex disease, and its pathogenesis depends on the interaction of a susceptible host with a multitude of environmental risk factors. The disease is characterized by disturbed alveologenesis. The many factors that influence alveologenesis include growth factors, cytokines and other substances that may act as ligands, receptors, signaling molecules and transcription factors, and the proteins that are the products of cell activity,

such as enzymes participating in matrix reconstruction, retinoids and elastin [[2–4](#page-9-0)].

Several experimental trials indicate that growth factors, especially those associated with vascularization (VEGF-Vascular endothelial growth factor), are closely related to the morphological changes in the respiratory tract of children with BPD [[3,5–8](#page-9-0)]. Researchers are also investigating inflammatory mediators, such as Tumor Necrosis Factor α (TNF- α), Interleukin-1β (IL-1β) Interleukin-6 (IL-6), Interleukin-8 (IL-8) and Interleukin-10 (IL-10) [[9–11\]](#page-9-0). Atypical pathogens, especially *Ureaplasma* spp., are believed to play a particular role in the inflammatory reaction leading to BPD [\[12,13\]](#page-9-0).

It is generally agreed that respiratory support in VLBW infants must be conducted in such a way as to circumvent damage caused by pressure, volume or oxygen. Results of a meta-analysis conducted by Stevens et al. demonstrate that extubation after early surfactant therapy and subsequent respiratory assistance with nasal continuous positive airway pressure results in a lower incidence of BPD compared with selective surfactant therapy and subsequent mechanical ventilation [\[14\]](#page-9-0). Other authors have presented similar observations favoring less invasive methods of respiratory assistance and lower ventilation values [[15](#page-9-0),[16](#page-9-0)]. Oxygen therapy and the subsequent action of its derivates (free radicals) has been proven to increase the incidence of BPD [[17](#page-9-0),[18](#page-9-0)]. To prevent such complications, practical guidelines recommending lower blood oxygen saturation values for preterm babies have been introduced [\[19,20](#page-10-0)].

Genetic foundations for the development of BPD are implicated in twin studies, which reveal highly significant concordance rates for BPD: 3.69-fold in monozygotic and 1.4 fold in dizygotic twins [[21](#page-10-0)]. Introduction of the microarray technique into clinical studies was one of the most important breakthroughs responsible for the dramatic progress in the field of human genetics during the last decade. The use of microarrays has given a new opportunity for studying even 20 000 human genes in a single experiment. The scope of potential applications of the microarrays is very broad, combining research and clinical medicine. The greatest advantage of this method is that it enables assessment of a great number of genetic factors (practically all human gene expression) although only a small amount of blood is necessary for testing, which is very important in preterm neonates.

To explore the pathogenesis of BPD, we carried out genome wide transcriptional profiling of RNA extracted from peripheral blood mononuclear cells of BPD subjects and non-BPD controls followed by pathway enrichment analysis in an attempt to identify biological pathways that were preferentially associated with BPD.

Methods

A prospective study was conducted between September 1, 2008 and November 30, 2010. The entry criteria were (a) preterm birth < 32 weeks gestational age, (b) birthweight ≤1500g, (c) the need for respiratory support. All patients were outborn in local hospitals and transported to the Polish American Children's Hospital which is a tertiary care unit for the region.

The majority of patients are referred from first-level neonatal care hospitals, which provide mainly for rural areas. Detailed perinatal history (birth weight, gestational age, Apgar score at 1 and 5 minutes after birth) and history of treatment in the referral hospital (mechanical ventilation, oxygen therapy, surfactant treatment, diagnoses) were taken on admission. Maternal fever/infection was used as surrogate for "clinical chorioamnionitis". Data on histological chorioamnionitis were unavailable in most cases. Ureaplasma infection was defined as positive tracheal aspirate culture for *Ureaplasma*

urealythicum. The study was approved by the Ethics Committee of the Jagiellonian University, Faculty of Medicine (KBET 27/B/2007). Written informed consent was obtained from all participants involved in the study.

Microarray analysis

After obtaining written informed consent from the parents, blood samples (0.3 ml) were drawn from all the study participants on the 5th, 14th and 28th day of life (DOL) for the assessment of whole genome expression in peripheral blood leukocytes. Subsequently, Ficoll isopaque gradient centrifugation (30 min, 2100 rpm, RT), and two times wash in 1x PBS (12 min, 1600 rpm, 40C), and finally RiboPure Blood Kit (Ambion, Life Technologies, Carlsbad, USA) were used for total RNA extraction.RNA concentration was measured with the use of NanoDrop Spectrophotometer (NanoDrop ND-1000; Thermoscientific, Waltham, USA), and RNA quality was determined by 2100 Bioanalyzer (Agilent, Santa Clara, USA).

100ng of total RNA were used for the single microarray experiment. GeneChip Human Gene 1.0 ST Arrays (Affymetrix, Santa Clara, USA) were used. Whole transcript microarray experiment was performed according to the manufacturer's protocol (GeneChip Whole Transcript sense Target Labeling Assay Manual, Version 4). The Affymetrix GeneChip Whole Transcript sense Target Labeling Assay is designed to generate amplified and biotinylated sense-strand DNA targets from the entire expressed genome. The protocol is optimized for use with the GeneChip Sense Target Arrays, where the probes are distributed throughout the entire length of each transcript. Details of microarray experiment were presented in online supplemental file – [Appendix S1.](#page-9-0)

Real-time PCR validation

To validate the results obtained by the microarray analysis real-time PCR technique was used. A total of 30 cDNA samples (15 patients with BPD and 15 controls) remaining after microarray analysis, were used for the validation procedure. Samples were randomly selected from the studied groups.

From each sample 100ng of cDNA was used for the single TaqMan Gene Expression Assay. Amplification reaction was performed with the use of TaqMan Universal PCR Master Mix and appropriate TaqMan probes (Life Technologies, Carlsbad, USA). Each sample was analyzed in duplicate.

Average between the expression of endogenous controls (housekeeping genes: GAPDH and Actin-B) was used for determination of the relative expression levels with the use of∆∆Ct calculation.For the validation procedure 14 genes were randomly selected from the group of genes differentially expressed in the microarray experiment (fold change greater or equal to $+/-1.5$) and from the genes without significant differences in expression.

Following TaqMan Gene Expression Assays were used: Hs00900055_m1 (VEGF gene), Hs00952786_m1 (AK5 gene), Hs00217864_m1 (OLAH gene), Hs01030384_m1 (ILR2 gene), Hs00187022_m1 (ADAM23 gene), Hs00736937_m1 (DAAM2 gene), Hs00360669_m1 (CD177 gene), Hs00255338_m1 (KLRC4 gene), Hs00171191_m1 (FBN1 gene), Hs00539582_s1 (LRRN3 gene), Hs00196254_m1 (NELL2

gene), Hs00924296_m1 (MPO gene), Hs00541549_m1 (ABCA13 gene), Hs00197437_m1 (OLFM4 gene),GAPDH (Hs02758991_g1) and Actin-B (Hs99999903_m1); (Life Technologies, Carlsbad, USA).

Outcome variables

Scheduled examinations were performed on all patients. The procedures of oxygen therapy in the examined group were consistent throughout the study with the goal of maintaining hemoglobin saturation between 88 and 95%. BPD severity was graded according to criteria proposed by Jobe and Bancalari [[22](#page-10-0)]. Four cohorts of children were identified based on examination at age 36 weeks post conception or on discharge. Mild BPD was defined as a need for supplemental oxygen for > or =28 days but not at 36 weeks' postmenstrual age (PMA) or discharge; moderate BPD as a need for supplemental oxygen for $>$ or =28 days plus treatment with <30% oxygen at 36 weeks' PMA, and severe BPD as a need for supplemental oxygen for $>$ or =28 days plus $>$ or =30% oxygen supplementation and/or positive pressure at 36 weeks' PMA. Infants treated with oxygen > 21% and/or positive pressure for nonrespiratory disease (e.g., central apnea or diaphragmatic paralysis) were not included in the BPD group unless they also developed parenchymal lung disease and exhibited clinical features of respiratory distress. Treatment with oxygen > 21% and/or positive pressure at 36-week postmenstrual age (PMA) due to an "acute" event was not a criterion for BPD. The evaluation of oxygen requirement was performed by a staff member blinded to molecular studies results.

Statistical analysis

The R package, *sizepower* was used to calculate sample size [[23](#page-10-0)]. A completely randomized treatment-control design was chosen. The total number of probe sets (genes) on the single microarray was 32323 (GeneChip Human Gene 1.0 ST Arrays – Affymetrix) and the number of genes that were not differentially expressed but falsely declared by the test procedure to be differentially expressed (number of false positives) was assumed to be about 1%. To simplify calculations, normal distribution of normalized gene expression values, the equality of variances and no interaction among genes were assumed. Based on preliminary data we anticipated that the variance of log-transformed gene expression levels should be equal to 0.6. The inclusion of 94 newborns (47 in each group) had power equal to 90% for detecting differentially expressed genes if significant gene expression difference was defined as fold expression change equal or greater than 1.5.

Basic demographic data were compared using the one-way analysis of variance or Kruskal-Wallis analysis of variance as appropriate. Qualitative values were compared using the chisquare test.

Neonatal data used for the study was recorded daily during hospitalization in NICU in a prospective manner and stored in computer databases. For the purpose of the study the following data was used: sex, birthweight, gestational age, intrauterine growth parameters, Apgar score, incidence of preeclampsia, maternal diabetes, preterm rupture of membranes, chorionamnionitis, delivery type, delivery room management, presence of respiratory distress syndrome (RDS), length of mechanical ventilation, surfactant administration, use of ibuprofen for patent ductus arteriosus (PDA), PDA ligation, early- and late-onset septic episodes, ureaplasma infection, prevelance of intraventricular haemorrhage (IVH), periventricular leukomalacia (PVL), weight gain during NICU stay and length of hospitalization.

The microarray data were pre-processed using the R/ Bioconductor package *aroma.affymetrix* [\[24,25\]](#page-10-0). For normalization we used Robust Multiarray Average (RMA) [\[26\]](#page-10-0). Quality control was performed by investigating Principal Component Analysis (PCA), *Relative Log Expression* (RLE) and Normalized Unscaled Standard Error (NUSE) plots. None of the arrays were discarded due to poor quality. Details of quality control analysis were presented in the online supplemental file – Figure S1-S7 in [Appendix S1](#page-9-0)

Moderated t-tests [[27](#page-10-0)] were performed to detect the probes that were differentially expressed between groups, using the *limma* package [\[28\]](#page-10-0) in the R statistical software. Multiple testing correction, using the Benjamini-Hochberg procedure was applied to control the false discovery rate (FDR) [\[29\]](#page-10-0). If a probe set had an corrected for multiple comparison p-value less than 0.01, it was declared significantly differentially expressed.

The groups differed with respect to birthweight, gestational age, delivery type, gender, Apgar score, delivery room intubation, surfactant therapy, prenatal steroids, use of ibuprofen for patent ductus arteriosus (PDA), PDA ligation, length of mechanical ventilation). The next step of the analysis was designed to examine whether the differences in gene expression between the groups were due to the differences between the groups with respect to these factors. A linear model was fitted to the expression data for each probe, and moderated *t*-tests and *F*-statistics were computed for each contrast. For each probe set that was found to be differentially expressed between the groups in the first part of the analysis, we tested whether this expression was explained by the group indicator and/or by gestational age and/or delivery type, gender, Apgar score, delivery room intubation, surfactant therapy, prenatal steroids, patent ductus arteriosus, length of mechanical ventilation using the function *classifyTestsF* from the limma package.

For the four categories of BPD we used Spearman's rho statistic [\[30\]](#page-10-0) to test whether there was a monotonic trend in gene expression between the categories.

DAVID annotation tools were used to explore which predefined gene sets were significantly enriched in one group compared to another [\[31,32](#page-10-0)]. The KEGG (Kyoto Encyclopedia of Genes and Genomes; www.genome.jp/kegg) and Biocarta pathways (www.biocarta.com) were selected for analysis. Genes differentially expressed between BPD patients and controls with p-values adjusted for multiple testing below 0.01 and with fold change greater than 1.5 were used as input for pathway enrichment analysis in DAVID.

The mean fold change values representing differences in expression between the BPD group and the control group with regard to each of the 14 genes selected for validation were compared between the microarray and the TaqMan Gene Expression experiments. The Student's *t*-test was used for evaluating the statistical significance of observed differences between gene expression.

Results

Characteristic of the studied group

One hundred twenty newborns were included in the study. The mean birth weight was 1029g (SD: 290), and the mean gestational age was 27.8 weeks (SD: 2.5). The majority of the pregnancies were terminated by abrupt deliveries. Only 50 (42%) of the mothers received antenatal steroids. Forty-nine (41%) were born by vaginal delivery. One hundred fourteen neonates (95%) required respiratory support. Seventy-three (61%) newborns received surfactant treatment. During the follow-up period, nine newborns died before the end of the study, therefore 111 children were included in the final analysis.

Bronchopulmonary dysplasia

During the follow-up period, BPD was diagnosed in 68 (61%) infants, including 40 (36%) children with mild disease, 13 (12%) with moderate and 15 (13%) with severe BPD. Forty-three newborns served as a control group (no BPD). The infants in the BPD group were of significantly lower birthweight and gestational age. As shown in Table 1 more infants in the BPD group received surfactant treatment and had persistent ductus arteriosus.

Whole genome expression

All data was collected and analyzed in the adherence to the Minimal Information About a Microarray Experiment guidelines. All primary microarray data was submitted to GEO public repository and are accessible through GEO Series accession number GSE32472 (http://www.ncbi.nlm.nih.gov/geo/query/ acc.cgi?acc=GSE32472).

A summary of the number of differentially expressed genes between the control and BPD patients is presented in [Table 2](#page-4-0). Overall 2086 genes were differentially expressed on the 5th DOL, only 324 on the 14th DOL, and 3498 on the 28th DOL. Generally, in each of the three time points during the first month of life more genes were under-expressed than overexpressed in the BPD group. The difference in expression measured as a fold change ranged between 1.0 and 1.5 in the majority of genes.

The results of the multivariate analysis are presented in the [Tables 3](#page-5-0) and [4.](#page-5-0) Among 2086 genes differentially expressed on the 5th day of life, only 255 were differentially expressed due to BPD alone. Most of the genes were differentially expressed due to the interaction of two factors: BPD and gestational age. Similar results were obtained by the analysis of genes differentially expressed on the $14th$ and $28th$ day of life. The observed differences in gene expression were moderate, only in a few cases the fold change values were higher than 1.5. The heatmaps from clustering analysis of genes with known gene symbol were presented in the online data supplement - Figures S8-S10 in [Appendix S1](#page-9-0).

Table 1. Comparison of selected demographic data in the bronchopulmonary dysplasia (BPD) groups and in the control children.

b_p-value for chi-square test

cp-value for Kruskal-Wallis analysis of variance

doi: 10.1371/journal.pone.0078585.t001

Tables S1, S2 and S3 in [Appendix S1](#page-9-0) present the one hundred genes with the highest difference in expression on the 5th, 14th and 28th DOL, respectively. Tables S4, S5 and S6 in [Appendix S1](#page-9-0) present results of multivariate analysis - the **Table 2.** Summary of the number of differentially expressed genes between bronchopulmonary dysplasia and control patients.

doi: 10.1371/journal.pone.0078585.t002

genes that were differentially expressed on the 5th, 14th and 28th DOL due to BPD alone. Among the over-expressed genes, there were some known factors contributing to the development of BPD such as growth factors and their receptors. In the group of under-expressed genes, the majority were involved in the inflammatory response.

Pathway enrichment analysis

Based on the results presented in Tables 2, [3](#page-5-0) and [4](#page-5-0), seven pathway enrichment analysis were performed. A summary of the analysis is presented in [Table 5](#page-6-0). The pathway up-regulated in the group of BPD patients is the cell cycle pathway. The activation of the cell cycle pathway does not seem to be related to the maturity of the infant. Detailed results of the cell cycle activation pathway are presented in Table S7 in [Appendix S1](#page-9-0).

Four pathways were continuously down-regulated on the $5th$, 14th and 28th day of life in the group of BPD patients. However, expression of the genes assigned to these pathways depended on both factors: immaturity and disease. Detailed results of the T cell receptor signaling pathways are presented in Table S8 in [Appendix S1.](#page-9-0)

Genes expression and BPD severity

To detect the presence of a monotonic trend in gene expression between the BPD severity groups Spearman's rho statistic was used. 1699 genes presented a monotone trend on the 5th, 1223 on the 14th and 4766 on the 28th day of life. Genes presenting a monotone trend in gene expression with p-values adjusted for multiple testing below 0.01 were used as input for pathway enrichment analysis in DAVID. Pathway enrichment analysis revealed similar results to those obtained in the two major groups (no BPD and BPD), see [Table 6](#page-7-0).

Real-time PCR validation

The microarray quality control was performed as recommended by Affymetrix. GeneChip Operating Software

built-in algorithms were used for this purpose. All the microarrays fulfilled the quality criteria. The validation procedure did not reveal significant differences between the results obtained with use of microarrays compare to real time PCR technique [Figure 1](#page-7-0) gives details on the above analysis.

Discussion

In this whole genome expression study of 5, 14 and 28 days old newborn children born before 32 weeks of gestation we have compared children who developed BPD to those without BPD. The nature of the study was mainly exploratory. We found a large number of differentially expressed genes between BPD and non BPD children. However, the difference in the expression measured as a fold change in the majority genes ranged between 1.0 and 1.5. The expression of only a few was twofold increased. Most of the genes were differentially expressed on the $5th$ and $28th$ day of life, and only a small number on the 14th day of life. More genes in the BPD group, were under-expressed rather than over-expressed. When adjusting for gender, gestational age and prenatal steroid treatment, the largest group of differentially expressed genes, were differentially expressed due to the combination of two factors: gestational age and BPD. All this general data are in accordance with our understanding of BPD. Recently described attempts to create logistic regression models directed towards BPD prediction conducted on large patient cohorts point to prematurity as the most important risk factor of this condition [[33](#page-10-0),[34](#page-10-0)].

Detailed analysis of the genes with the highest difference in expression was performed. Interestingly, in this group of genes we did not find many of the factors traditionally associated with the development of BPD (inflammatory mediators, growth factors).

A different method of evaluation of gene expression alterations is pathway enrichment analysis. This method can detect transcriptional programs that are distributed across an

Table 3. Number of differentially expressed genes between the BPD and control group – results of multivariate analyses.

doi: 10.1371/journal.pone.0078585.t003

entire network of genes yet are subtle at the level of individual genes. Based on this method we found two groups of pathways

Table 4. Number of differentially expressed genes between the BPD and control group – results of multivariate analyses.

. included in pathway analysis 7

doi: 10.1371/journal.pone.0078585.t004

which were differentially regulated at day 5, 14 and 28 after birth in infants with BPD. The first is the cell cycle pathway which seems to be a maturity independent pathway significantly up-regulated in children with BPD. The second group included pathways involved in inflammatory response. These pathways were down-regulated in children with BPD. The expression of genes involved in these pathways was related to both maturity and the severity of BPD.

Cell cycle

Pathway enrichment analysis, based on gene expression fold change indicated the cell cycle pathway as one of the most up-regulated pathways among the BPD newborns. The genes assigned to this pathway with the highest fold change values are genes encoding the family of cyclins (*CYCA, CYCE* and *CYCB*) and cyclin related kinase 1 (*CDK1*). There was no difference in the expression of *CYCD* and *CDK4/CDK6*, as well as *CDK2* as a partner for cyclin E and A. This situation can be analyzed from several different aspects. The up-regulation of the cell cycle pathway might be considered an indicator of an increased cell mitosis rate and increased cell hyperplasia.

In adult and newborn mice hyperoxia induced expression of p21 protein responsible for the inhibition of cyclin E dependent kinase activity. Moreover, results from some earlier animal experiments indicate that hyperoxia inhibits cell proliferation by two mechanisms; direct cell toxicity and modulation of G1/S phase transition [\[35\]](#page-10-0).

The precise interpretation of our results is difficult in view of the highly specific character of the cell cycle and the still elusive knowledge on the gene interactions taking part in these complicated processes. The majority of cell cycle genes upregulated in the BPD newborns are involved in the regulation of G1/S transition (*CYCE*), S/G² transition (*CYCA* and *CDK1*), and G2/M transition (*CYCB* and *CDK1*). The question can be raised, however, why neither *CYCD* nor *CDK4* and *CDK6* showed overexpression. S/G² transition is under control of cyclin A and cyclin depended kinases CDK1 and CDK2. In our studied group overexpression of only *CYCA* and *CDK1* was observed. However, based on the experimental studies in CDK2 knockout mice [[36\]](#page-10-0) it was suggested that CDK2 is not required in the cell cycle progress and might be substituted by

Table 5. Summary of the seven pathway analysis for the differentially expressed genes between the BPD and control group.

doi: 10.1371/journal.pone.0078585.t005

CDK1. Finally, the G2-M transition is controlled by the *CYCB/ CDK1* complex whose increased expression was observed in the BPD group.

As the level of cyclin B increases, it is bound to CDK1 and then transported to the nucleus promoting the beginning of the M phase. This process can be stopped by the phosphorylation

Table 6. Pathway enrichment analysis of the genes presented a monotone trend in the gene expression between the BPD severity groups.

Pathways with FDR value less than 10% are shown.

doi: 10.1371/journal.pone.0078585.t006

of CDK. To protect CDK against this phosphorylation, a specific phosphatase cdc25 and Wee kinase are needed. In fact both of the above mentioned genes (*CDC25* and *WEE*) were found to be over-expressed in the BPD group. These results are therefore strongly indicative of increased mitotic activity of leukocytes originating from newborns who developed BPD. The *P53* gene representing the major "sensor" of DNA damage was under-expressed in the BPD group compared to the controls, and all genes coding for cyclin kinase inhibitory proteins like *P21*, *P27*, and *P57* showed comparable expression to those of the controls. This data might suggest that the group of BPD newborns was not exposed to, or was better protected against the influence of environmental factors potentially leading to DNA damage.

However, it should be noted that the used microarray approach could reveal non-specific cell cycle pathway upregulation. Some studies showed that there is a striking similarity between the gene expression responses to *a priori* very different phenomena. Abrupt temperature changes, oxidative stress or the addition of particular molecules are examples of such perturbations [[37](#page-10-0),[38](#page-10-0)]. A change in the amount of nutrients is another example [\[39\]](#page-10-0).

In summary, these data suggest strong cell mitotic activity in the group of children who developed BPD compared to the controls, with weaker indications of activation of DNA repair mechanisms.

T cell receptor signaling pathway

T cells are activated upon the triggering of T cell receptors by antigen presenting cells. T cell activation induces a series of biochemical events that can differentially signal the naive T cell to enter into a pathway leading to the generation of effector T cells with the onset of rapid proliferation and production of effector cytokines; or enter into a state of antigenic nonresponsiveness known as anergy or die by apoptosis [[40](#page-10-0)]. The T cell response depends on the type of ligand bound to its' receptors, time of co-operation, presence of co-receptors or coinhibitors [\[41,42\]](#page-10-0).

The production of cytokines is essentially controlled by transcription factors [\[43\]](#page-10-0). In our study, transcription factors and related pathways were under-expressed in children with BPD. When the differential expression analysis was adjusted for gender, gestational age and prenatal steroid treatment, it was revealed that not only immaturity is responsible for the downregulation of genes in this pathway, but also disease severity.

The significance of the finding decreased inflammatory genes expression might be unknown. However, we can speculate that decreased expression of T cell receptors may lead to relative anergy and can be a risk factor for bacterial translocation and further infection. These findings correspond with our knowledge that one of the risk factors of BPD is pulmonary infection [[44](#page-10-0)]. Also, there are data indicating that expression of genes involved in the T cell receptors signaling pathway can be decreased by reactive oxygen species [\[45\]](#page-10-0).

Analysis of single gene expression of the genes involved in these pathways revealed that some genes from this signaling pathway were over-expressed in BPD children. All the overexpressed genes in this signaling pathway belong to a classical

part of the mitogen-activated protein kinase signaling pathway. These genes can be activated by growth factors or in response to pro-inflammatory stimuli. This signaling pathway is involved not only in the differentiation or proliferation of lymphocytes, but also of other cell lines.

Limitations of the study

This type of study protocol raises the question of whether gene expression of white blood cells reflects gene expression in the lung. Investigators collaborating on the MuTHER project (Multiple Tissue Human Expression Resource) have addressed this issue by studying gene expression in different tissues in the same individuals [\[46\]](#page-10-0). The researchers collected skin, subcutaneous fat and peripheral blood from female twins. They found that overall 50 to 80% the loci identified in one tissue were estimated to have gene regulatory effects in a comparison tissue. Another interesting study was conducted by Rudkowska et al. [\[47\]](#page-10-0). The objective of this study was to validate the use of peripheral blood mononuclear cells for gene expression analysis as a marker of nutritional intervention as an alternative to skeletal muscle tissue biopsies. The author found that 88% (32,341) of transcripts were coexpressed in both tissues. Importantly, very strong correlation was observed between transcript expression levels of peripheral blood mononuclear cells and skeletal muscle tissue biopsies.

There is no data about the relationship between gene expression analysis in circulating cells and lung tissue, however we can only speculate, on the grounds that hypoxia/ hyperoxia influences the organism of the newborn as a whole, that analysis of gene expression within the peripheral blood leukocytes which are characterized by high mitotic activity, may reflects gene expression in other cells.

Of note is that the neonatal unit where the study was conducted accepts a highly selective group of newborns from primary and secondary referral centers. Thus, the study focused on a population with a high risk for BPD development. The specific character of the studied population might have influenced the results that were obtained; namely they were biased towards a statistically significant difference. The patient population may not be applicable to other NICUs (for example: the low rate of prenatal steroid use), and hence our results may not be generalizable. Another problem relates to the results of the statistical analysis in the case of multiple comparisons. If no correction was applied, multiple testing would increase the probability of a type I error to 1.0 (100%). Bearing this fact in mind, the correction for the multiple tests was performed, however at least 1% of false discovery rate can still be expected. The estimated power of the study to reveal a 1.5 fold change difference between the BPD and no BPD groups was 92.8%.

It should be noted that the genome expression profiles detected in our study may be confounded by limitations related to the nature of biological systems which are dynamic and change constantly. The mRNAs are being synthesized and degraded at its own rate and the gene expression is measured during the experiment at a steady state. Thus, the presence of mRNA does not necessarily mean that it was just synthesized and the inability to detect certain transcripts may be due to their

quick degradation. In addition, many biologically important changes do not manifest themselves in alterations of the RNA levels. Most of the cellular functions are performed by proteins and alteration to these functions can result not only from changes in protein levels but also from protein modifications [[48](#page-10-0)].

Conclusions

The results of this whole genome expression study revealed alteration of the expression of nearly 10% of the genome in BPD patients. Most of the differences in expression were associated with the degree of immaturity of the infant. However, based on pathway enrichment analysis we found pathways potentially involved in the pathogenesis of the disease. These results can be considered in two aspects. One, by reflecting our current knowledge and two, by providing some new information on the genes potentially involved in the BPD patomechanism. The results indicating on the overexpression of the pathways involving cytokine and theirs receptors, corroborates the generally accepted concept of the contribution of inflammatory response to the etiology of BPD. On other hand however, we observed lower expression of different immune response pathways, including T cell receptor pathways. These may point out on the relative immunedeficiency due to the extreme immaturity of the newborn. The second aspect of our results indicates some unknown facts, which may throw a new light on the BPD

References

- 1. Bhandari A, Bhandari V (2009) Pitfalls, problems, and progress in bronchopulmonary dysplasia. Pediatrics 123: 1562-1573. doi[:10.1542/](http://dx.doi.org/10.1542/peds.2008-1962) [peds.2008-1962](http://dx.doi.org/10.1542/peds.2008-1962). PubMed: [19482769](http://www.ncbi.nlm.nih.gov/pubmed/19482769).
- 2. Warburton D, Bellusci S, De Langhe S, Del Moral PM, Fleury V et al. (2005) Molecular mechanisms of early lung specification and branching morphogenesis. Pediatr Res 57: 26R–37R. doi:[10.1203/01.PDR.](http://dx.doi.org/10.1203/01.PDR.0000159570.01327.ED) [0000159570.01327.ED.](http://dx.doi.org/10.1203/01.PDR.0000159570.01327.ED) PubMed: [15817505.](http://www.ncbi.nlm.nih.gov/pubmed/15817505)
- 3. Thébaud B (2007) Angiogenesis in lung development, injury and repair: implications for chronic lung disease of prematurity. Neonatology 91: 291–297. doi[:10.1159/000101344](http://dx.doi.org/10.1159/000101344). PubMed: [17575472](http://www.ncbi.nlm.nih.gov/pubmed/17575472).
- 4. Bourbon J, Boucherat O, Chailley-Heu B, Delacourt C (2005) Control mechanisms of lung alveolar development and their disorders in bronchopulmonary dysplasia. Pediatr Res 57: 38R–46R. doi: [10.1203/01.PDR.0000159630.35883.BE.](http://dx.doi.org/10.1203/01.PDR.0000159630.35883.BE) PubMed: [15817499.](http://www.ncbi.nlm.nih.gov/pubmed/15817499)
- 5. Asikainen TM, Ahmad A, Schneider BK, White CW (2005) Effect of preterm birth on hypoxia-inducible factors and vascular endothelial growth factor in primate lungs. Pediatr Pulmonol 40: 538–546. doi: [10.1002/ppul.20321](http://dx.doi.org/10.1002/ppul.20321). PubMed: [16231377](http://www.ncbi.nlm.nih.gov/pubmed/16231377).
- 6. Janér J, Lassus P, Haglund C, Paavonen K, Alitalo K et al. (2006) Pulmonary vascular endothelial growth factor-C in development and lung injury in preterm infants. Am J Respir Crit Care Med 174: 326– 330. doi[:10.1164/rccm.200508-1291OC.](http://dx.doi.org/10.1164/rccm.200508-1291OC) PubMed: [16690974.](http://www.ncbi.nlm.nih.gov/pubmed/16690974)
- 7. Lassus P, Turanlahti M, Heikkilä P, Andersson LC, Nupponen I et al. (2001) Pulmonary vascular endothelial growth factor and Flt-1 in fetuses, in acute and chronic lung disease, and in persistent pulmonary hypertension of the newborn. Am J Respir Crit Care Med 164: 1981– 1987. doi[:10.1164/ajrccm.164.10.2012036.](http://dx.doi.org/10.1164/ajrccm.164.10.2012036) PubMed: [11734455.](http://www.ncbi.nlm.nih.gov/pubmed/11734455)
- 8. Thébaud B, Abman SH (2007) Bronchopulmonary dysplasia: where have all the vessels gone? Roles of angiogenic growth factors in chronic lung disease. Am J Respir Crit Care Med 175: 978–985. doi: [10.1164/rccm.200611-1660PP.](http://dx.doi.org/10.1164/rccm.200611-1660PP) PubMed: [17272782.](http://www.ncbi.nlm.nih.gov/pubmed/17272782)
- 9. Bry K, Lappalainen U (2006) Pathogenesis of bronchopulmonary dysplasia: the role of interleukin 1b in the regulation of inflammationmediated pulmonary retinoic acid pathways in transgenic mice. Semin Perinatol 30: 121–128. doi:[10.1053/j.semperi.2006.04.009](http://dx.doi.org/10.1053/j.semperi.2006.04.009). PubMed: [16813970.](http://www.ncbi.nlm.nih.gov/pubmed/16813970)

patomechanism. This refers to the significant overexpression of the cell cycle pathway in BPD patients. These findings should be considered with caution for two reasons. First, the over- and under-expression of the genes are dynamic process strictly dependent on the phase of the cell cycle, therefore the assessment performed at the single time point might be the resultant of different phases of the examined cells. Second, the final outcome represented by turning genes on or off is the result of concerted action of many other genes, mainly transcription factors, which activity might last too short to be determined. The above reservations do not change the fact, however, that the results are of sufficient interest to warrant further investigation on this subject.

Supporting Information

Appendix S1. Detailed description of microarray evaluation, tables S1-S8, figures S1-S10. (DOCX)

Author Contributions

Conceived and designed the experiments: JJP PK EJW ODS. Performed the experiments: PK MJ TT. Analyzed the data: PK CCG. Contributed reagents/materials/analysis tools: MBM AMT. Wrote the manuscript: JJP PK ODS.

- 10. Nedrelow JH, Bhandari V (2004) Interleukin (IL)-6 to vascular endothelial growth factor (VEGF) ratio predicts the development of brochopulmonary dysplasia (BPD)/death in premature infants. Pediatr Res 55: 496A–496A.
- 11. Vento G, Capoluongo E, Matassa PG, Concolino P, Vendettuoli V et al. (2006) Serum levels of seven cytokines in premature ventilated newborns: correlations with old and new forms of bronchopulmonary dysplasia. Intensive Care Med 32: 723–730. doi[:10.1007/](http://dx.doi.org/10.1007/s00134-006-0138-1) [s00134-006-0138-1.](http://dx.doi.org/10.1007/s00134-006-0138-1) PubMed: [16550369.](http://www.ncbi.nlm.nih.gov/pubmed/16550369)
- 12. Schelonka RL, Waites KB (2007) Ureaplasma infection and neonatal lung disease. Semin Perinatol 31: 2–9. doi:[10.1053/j.semperi.](http://dx.doi.org/10.1053/j.semperi.2007.01.001) [2007.01.001](http://dx.doi.org/10.1053/j.semperi.2007.01.001). PubMed: [17317421](http://www.ncbi.nlm.nih.gov/pubmed/17317421).
- 13. Colaizy TT, Morris CD, Lapidus J, Sklar RS, Pillers DA (2007) Detection of ureaplasma DNA in endotracheal samples is associated with bronchopulmonary dysplasia after adjustment for multiple risk factors. Pediatr Res 61: 578–583. doi:[10.1203/pdr.0b013e318045be03](http://dx.doi.org/10.1203/pdr.0b013e318045be03). PubMed: [17413863.](http://www.ncbi.nlm.nih.gov/pubmed/17413863)
- 14. Stevens T, Harrington E, Blennow M, Soll R (2007) Early surfactant administration with brief ventilation vs. selective surfactant and continued mechanical ventilation for preterm infants with or at risk for respiratory distress syndrome. Cochrane Database Syst Rev 4, CD003063.
- 15. Kamlin C, Davis PG (2004) Long versus short inspiratory times in neonates receiving mechanical ventilation. Cochrane Database Syst Rev 4, CD004503. PubMed: [15495117](http://www.ncbi.nlm.nih.gov/pubmed/15495117).
- 16. Cools F, Henderson-Smart D, Offringa M, Askie L (2009) Elective high frequency oscillatory ventilation versus conventional ventilation for acute pulmonary dysfunction in preterm infants. Cochrane Database
Syst Rev 3, CD000104. PubMed: Syst Rev 3, CD000104. PubMed: [195883171763659014583909125353831168695010796129](http://www.ncbi.nlm.nih.gov/pubmed/195883171763659014583909125353831168695010796129).
- 17. Saugstad OD (2010) Oxygen and oxidative stress in bronchopulmonary dysplasia. J Perinat Med 38: 571-577. PubMed: [20807008](http://www.ncbi.nlm.nih.gov/pubmed/20807008).
- 18. Saugstad OD, Aune D (2011) In search of the optimal oxygen saturation for extremely low birth weight infants: a systematic review and meta-analysis. Neonatology 100: 1-8. doi[:10.1159/000322001](http://dx.doi.org/10.1159/000322001). PubMed: [21150224.](http://www.ncbi.nlm.nih.gov/pubmed/21150224)
- 19. Askie L, Henderson-Smart D, Ko H (2009) Restricted versus liberal oxygen exposure for preventing morbidity and mortality in preterm or low birth weight infants. Cochrane Database Syst Rev 1, CD001077.
- 20. Stevens TP, Dylag A, Panthagani I, Pryhuber G, Halterman J (2010) Effect of cumulative oxygen exposure on respiratory symptoms during infancy among VLBW infants without bronchopulmonary dysplasia. Pediatr Pulmonol 45: 371–379. PubMed: [20232470.](http://www.ncbi.nlm.nih.gov/pubmed/20232470)
- 21. Bhandari V, Bizzarro MJ, Shetty A, Zhong X, Page GP et al. (2006) Familial and genetic susceptibility to major neonatal morbidities in preterm twins. Pediatrics 117: 1901–1906. doi[:10.1542/peds.](http://dx.doi.org/10.1542/peds.2005-1414) [2005-1414](http://dx.doi.org/10.1542/peds.2005-1414). PubMed: [16740829](http://www.ncbi.nlm.nih.gov/pubmed/16740829).
- 22. Jobe AH, Bancalari E (2001) Bronchopulmonary dysplasia. Am J Respir Crit Care Med 163: 1723-1729. doi:[10.1164/ajrccm.](http://dx.doi.org/10.1164/ajrccm.163.7.2011060) [163.7.2011060](http://dx.doi.org/10.1164/ajrccm.163.7.2011060). PubMed: [11401896](http://www.ncbi.nlm.nih.gov/pubmed/11401896).
- 23. Lee M-T (2004) Power and Sample Size Considerations. In: Analysis of Microarray Expression Data. Boston: Kluwer Academic Publishers. pp. 191-236.
- 24. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80. doi[:10.1186/](http://dx.doi.org/10.1186/gb-2004-5-10-r80) [gb-2004-5-10-r80](http://dx.doi.org/10.1186/gb-2004-5-10-r80). PubMed: [15461798](http://www.ncbi.nlm.nih.gov/pubmed/15461798).
- 25. Bengtsson H, Simpson K, Bullard J, Hansen K (2008) aroma.affymetrix: A generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory. Tech Rep #745, Department of Statistics, University of California, Berkeley.
- 26. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249-264. doi: [10.1093/biostatistics/4.2.249](http://dx.doi.org/10.1093/biostatistics/4.2.249). PubMed: [12925520](http://www.ncbi.nlm.nih.gov/pubmed/12925520).
- 27. Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article3, Article3 PubMed: [16646809.](http://www.ncbi.nlm.nih.gov/pubmed/16646809)
- 28. Smyth GK (2005) Limma: Linear Models for Microarray Data. In: R GentlemanV CareyS DudoitR IrizarryW Huber. Bioinformatics and Computional Biology Solutions using R and Bioconductor. New York: Springer Verlag. pp. 397-420.
- 29. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B Stat Methodol 57: 289-300.
- 30. Hollander M, Wolfe DA (1973) Kendall and Spearman Tests. In: Nonparametric Statistical Methods. New York: John Wiley & Sons. pp. 185-194.
- 31. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57. PubMed: [19131956.](http://www.ncbi.nlm.nih.gov/pubmed/19131956)
- 32. Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: 1-13. doi[:10.1093/nar/](http://dx.doi.org/10.1093/nar/gkp505) [gkp505.](http://dx.doi.org/10.1093/nar/gkp505) PubMed: [19033363.](http://www.ncbi.nlm.nih.gov/pubmed/19033363)
- 33. Bhering CA, Mochdece CC, Moreira ME, Rocco JR, Sant'Anna GM (2007) Bronchopulmonary dysplasia prediction model for 7-day-old

infants. J Pediatr (Rio, 83: 163–70. PubMed: [17380230](http://www.ncbi.nlm.nih.gov/pubmed/17380230) J) 83: 163-170. . PubMed: 17380230

- 34. Henderson-Smart DJ, Hutchinson JL, Donoghue DA, Evans NJ, Simpson JM et al. (2006) Prenatal predictors of chronic lung disease in very preterm infants. Arch Dis Child Fetal Neonatal Ed 91: F40-F45. PubMed: 16131530
- 35. O'Reilly MA (2001) DNA damage and cell cycle checkpoints in hyperoxic lung injury: braking to facilitate repair. Am J Physiol Lung Cell Mol Physiol 281: L291-L305. PubMed: [11435201](http://www.ncbi.nlm.nih.gov/pubmed/11435201).
- 36. Sánchez I, Dynlacht BD (2005) New insights into cyclins, CDKs, and cell cycle control. Semin Cell Dev Biol 16: 311-321. doi[:10.1016/](http://dx.doi.org/10.1016/j.semcdb.2005.02.007) [j.semcdb.2005.02.007](http://dx.doi.org/10.1016/j.semcdb.2005.02.007). PubMed: [15840440](http://www.ncbi.nlm.nih.gov/pubmed/15840440).
- 37. Lepock JR (2005) How do cells respond to their thermal environment? Int J Hyperthermia 21: 681-687. doi[:10.1080/02656730500307298](http://dx.doi.org/10.1080/02656730500307298). PubMed: [16338849.](http://www.ncbi.nlm.nih.gov/pubmed/16338849)
- 38. Kriebitzsch C, Verlinden L, Eelen G, Tan BK, Van Camp M et al. (2009) The impact of 1,25(OH)2D3 and its structural analogs on gene expression in cancer cells--a microarray approach. Anticancer Res 29: 3471-3483. PubMed: [19667141.](http://www.ncbi.nlm.nih.gov/pubmed/19667141)
- 39. Lapillonne A, Clarke SD, Heird WC (2004) Polyunsaturated fatty acids and gene expression. Curr Opin Clin Nutr Metab Care. 7: 151-156. doi: [10.1097/00075197-200403000-00008](http://dx.doi.org/10.1097/00075197-200403000-00008). PubMed: [15075705](http://www.ncbi.nlm.nih.gov/pubmed/15075705).
- 40. Huang Y, Wange RL (2004) T Cell Receptor Signaling: Beyond Complex Complexes. J Biol Chem 279: 28827-28830. doi[:10.1074/](http://dx.doi.org/10.1074/jbc.R400012200) [jbc.R400012200](http://dx.doi.org/10.1074/jbc.R400012200). PubMed: [15084594](http://www.ncbi.nlm.nih.gov/pubmed/15084594).
- 41. Cheng J, Montecalvo A, Kane LP (2011) Regulation of NF-kappaB induction by TCR/CD28. Immunol Res 50: 113-117. doi[:10.1007/](http://dx.doi.org/10.1007/s12026-011-8216-z) [s12026-011-8216-z](http://dx.doi.org/10.1007/s12026-011-8216-z). PubMed: [21717079](http://www.ncbi.nlm.nih.gov/pubmed/21717079).
- 42. James JR, McColl J, Oliveira MI, Dunne PD, Huang E et al. (2011) The T-cell receptor triggering apparatus is composed of monovalent or monomeric proteins. J Biol Chem. doi:[10.1074/jbc.M111.219212.](http://dx.doi.org/10.1074/jbc.M111.219212)
- 43. Beyer T, Busse M, Hristov K, Gurbiel S, Smida M et al. (2011) Integrating signals from the T-cell receptor and the interleukin-2 receptor. PLOS Comput Biol 7: e1002121. PubMed: [21829342.](http://www.ncbi.nlm.nih.gov/pubmed/21829342)
- 44. Speer CP (2006) Inflammation and bronchopulmonary dysplasia: a continuing story. Semin Fetal Neonatal Med 11: 354-362. doi[:10.1016/](http://dx.doi.org/10.1016/j.siny.2006.03.004) [j.siny.2006.03.004](http://dx.doi.org/10.1016/j.siny.2006.03.004). PubMed: [16702036](http://www.ncbi.nlm.nih.gov/pubmed/16702036).
- 45. Kaminski MM, Sauer SW, Klemke CD, Süss D, Okun JG et al. (2010) Mitochondrial Reactive Oxygen Species Control T Cell Activation by Regulating IL-2 and IL-4 Expression: Mechanism of Ciprofloxacin-Mediated Immunosuppression. J Immunol 184: 4827-4841. doi: [10.4049/jimmunol.0901662.](http://dx.doi.org/10.4049/jimmunol.0901662) PubMed: [20335530.](http://www.ncbi.nlm.nih.gov/pubmed/20335530)
- 46. Grundberg E, Small KS, Hedman AK, Nica AC, Builet A et al. (2012) Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet 44: 1084-1089. doi[:10.1038/ng.2394.](http://dx.doi.org/10.1038/ng.2394) PubMed: [22941192.](http://www.ncbi.nlm.nih.gov/pubmed/22941192)
- 47. Rudkowska I, Raymond C, Ponton A, Jacques H, Lavigne C et al. (2011) Validation of the use of peripheral blood mononuclear cells as surrogate model for skeletal muscle tissue in nutrigenomic studies. OMICS 5: 1-7. PubMed: [21194298](http://www.ncbi.nlm.nih.gov/pubmed/21194298).
- 48. Kim K, Zakharkin SO, Allison DB (2010) Expectations, validity, and reality in gene expression profiling. J Clin Epidemiol 63: 950–959. doi: [10.1016/j.jclinepi.2010.02.018.](http://dx.doi.org/10.1016/j.jclinepi.2010.02.018) PubMed: [20579843.](http://www.ncbi.nlm.nih.gov/pubmed/20579843)