


RESEARCH ARTICLE

Open Access



Exome sequencing identifies gene variants and networks associated with extreme respiratory outcomes following preterm birth

Aaron Hamvas^{1,9*}, Rui Feng², Yingtao Bi³, Fan Wang², Soumyaroop Bhattacharya⁴, Jared Mereness⁴, Madhurima Kaushal⁵, C Michael Cotten⁶, Philip L Ballard⁷, Thomas J Mariani^{4,8*}  and PROP Investigators

Abstract

Background: Previous studies have identified genetic variants associated with bronchopulmonary dysplasia (BPD) in extremely preterm infants. However, findings with genome-wide significance have been rare, and not replicated. We hypothesized that whole exome sequencing (WES) of premature subjects with extremely divergent phenotypic outcomes could facilitate the identification of genetic variants or gene networks contributing disease risk.

Results: The Prematurity and Respiratory Outcomes Program (PROP) recruited a cohort of > 765 extremely preterm infants for the identification of markers of respiratory morbidity. We completed WES on 146 PROP subjects (85 affected, 61 unaffected) representing extreme phenotypes of early respiratory morbidity. We tested for association between disease status and individual common variants, screened for rare variants exclusive to either affected or unaffected subjects, and tested the combined association of variants across gene loci. Pathway analysis was performed and disease-related expression patterns were assessed. Marginal association with BPD was observed for numerous common and rare variants. We identified 345 genes with variants unique to BPD-affected preterm subjects, and 292 genes with variants unique to our unaffected preterm subjects. Of these unique variants, 28 (19 in the affected cohort and 9 in unaffected cohort) replicate a prior WES study of BPD-associated variants. Pathway analysis of sets of variants, informed by disease-related gene expression, implicated protein kinase A, MAPK and Neuregulin/epidermal growth factor receptor signaling.

Conclusions: We identified novel genes and associated pathways that may play an important role in susceptibility/resilience for the development of lung disease in preterm infants.

Keywords: Prematurity and respiratory outcomes program (PROP), Bronchopulmonary dysplasia (BPD)

Background

Bronchopulmonary dysplasia (BPD) is the most significant pulmonary consequence of premature birth, affecting 30–60% of infants born before 28 weeks of gestation, depending on definition of the disease and the racial/ethnic population studied [1]. In contrast to most morbidities associated with extreme prematurity, the prevalence of BPD among premature infants has been

stable or slightly rising over the past decade [2, 3]. Furthermore, emerging data suggests that preterm birth, and a diagnosis of BPD, is associated with increased risk for chronic lung diseases that present in older children and adults. Concordance of respiratory outcomes in monozygotic twins suggest that genetic factors may play a substantial role in influencing these outcomes [4, 5]. Genome-wide association studies (GWAS) attempting to identify BPD candidate genes using variants with a population-based frequency of more than 1% have either failed to identify candidate genes or identified genes that have not been replicated on subsequent studies [6–12].

* Correspondence: ahamvas@luriechildrens.org; Tom_Mariani@urmc.rochester.edu

¹Department of Pediatrics, Northwestern University, Chicago, IL, USA

⁴Department of Pediatrics, University of Rochester, Rochester, NY, USA

Full list of author information is available at the end of the article



The GWAS variant detection platform assumes that previously identified, generally common, variants with a minor allele frequency > 1% in the population, either inform directly or tag functional variants in the genome. However, a large percentage of variants in individual genomes have population-based minor allele frequencies of less than 1%, necessitating approaches that can identify rare and even unique variants. Because BPD is a complex trait that results from developmental and environmental as well as genetic factors, no single approach is likely to identify the likely multiple genes contributing to risk for BPD. Reasoning that variants that would contribute to BPD would be under negative selection pressure, Li et al. performed exome sequencing and identified extremely rare variants and associated networks of genes that were unique to the cohort with BPD [13].

The Prematurity and Respiratory Outcomes Program (PROP, <http://clinicaltrials.gov/ct2/show/NCT01435187>) is a geographically diverse multicenter effort across the United States to identify biochemical, physiological and genetic biomarkers that are predictive of later pulmonary status in preterm infants born at < 29 weeks of gestation [14, 15]. PROP enrolled 765 preterm infants who survived to 36-weeks of age. As an initial step in identifying genes that might be contributing to adverse short-term outcomes in this population, and because we have access to high-resolution phenotype definition in this cohort, we performed exome sequencing on a subset of infants with BPD phenotypes at the extremes of disease severity. Similar approaches have been used in small cohorts to identify candidate genes and genetic modifiers for several conditions, including cystic fibrosis, chronic obstructive pulmonary disease, macular degeneration, and preterm birth [16–20]. We hypothesized that whole exome sequencing of subjects with extreme phenotypic outcomes would be more likely to enrich for common and rare deleterious and protective genetic variants that contribute to disease risk, and we then utilized several analytic approaches with differing assumptions to identify signals that may contribute to the risk for BPD.

Methods

Subject and phenotype selection: Subjects for the PROP were infants born at < 29 weeks of gestation who were followed prospectively through discharge from the NICU [1, 15]. Infants with congenital anomalies and those who were not anticipated to survive the immediate neonatal period were excluded. Saliva was collected in Oragene® kits (DNA Genotek, Inc., Ottawa, Ontario, Canada) and DNA was extracted, quantified, and stored at a central archive [15] for the PROP study. Studies at each of the participating sites were approved by the local ethics boards and all parents consented to collection and

sequencing of infant DNA and reporting of de-identified results.

We defined extreme phenotypes of respiratory function at 36 weeks post menstrual age (PMA) that included; 1) the most “unaffected” group who required supplemental oxygen for < 28 days and did not require any respiratory support at 36 weeks ($n = 61$), and 2) the most “affected” group, who required mechanical ventilation for > 14 days and required continuous respiratory support through 36 weeks ($n = 85$). This respiratory support included invasive (mechanical ventilation) and non-invasive (CPAP, high flow nasal cannula > 2 l/min flow) approaches, with or without supplemental oxygen (i.e. some infants required respiratory support but with FiO_2 0.21). For the purposes of this study, infants who died before 36 weeks’ PMA were excluded. The characteristics of the population are displayed in Table 1.

Sequencing and variant annotation

Exome sequencing was performed through the National Heart, Lung, and Blood Institute (NHLBI) Resequencing and Genotyping Service at the University of Washington. The quality was filtered according to GATK3.0 recommended filters; variants with any “QUAL \leq 50.0, ABHet > 0.75, HRun > 4.0, QD < 5.0, SB \geq 0.10, number of ALT \geq 1, MQ < 40.0, or FS > 60.0” were removed. QUAL is Phred-scaled probability that support existence of ALT, ABHet is the allele balance measure for heterozygosity, HRun is the largest contiguous homopolymer run, QD is coverage-normalized quality score, SB is strand bias, MQ is RMS mapping quality, and FS is Fisher’s strand bias. The sequence data were high quality, with > 60 \times coverage (total sequence reads of all samples at each variant locus without filtering) and a transition to transversion ratio (Ti/Tv) of > 3, allowing the identification of 189,404 high confidence variants in the population, of which 149,344 (77.1%) had a minor allele frequency (MAF) < 0.05 (Additional file 1: Table S1). We used ANNOVAR (2015Mar22 version) to annotate all identified variants and functionally classified the exonic variants into categories including synonymous, nonsynonymous, stop gain/loss, splice site gain/loss, etc. [21]. In total, we identified 88,183 high-confidence common and rare nonsynonymous and otherwise presumed functional variants among the 146 individuals.

Covariate determination

To derive population substructures, we used the multidimensional scaling (MDS) method implemented in PLINK (<http://zzz.bwh.harvard.edu/plink/>, [22]) using all variants that overlap with 1000 genome data. The first two principal components from MDS were included as covariates in subsequent association models to adjust for population stratification bias. As multiple births accounted for 22

Table 1 Demographics

Infant Characteristics	Cohort total	Cases	Controls	Comparison ^a
Sample Size	146	85	61	
Gestational Age, days,	185 ± 10	180 ± 9	190 ± 8	< 0.001
Gestational Age, n (%), completed wks.				
	23	5 (3.4)	5 (5.9)	0 (0.0)
	24	23 (15.8)	22 (25.9)	1 (1.6)
	25	24 (16.4)	16 (18.8)	8 (13.1)
	26	36 (24.7)	23 (27.1)	13 (21.3)
	27	34 (23.3)	14 (16.5)	20 (32.8)
	28	24 (16.4)	5 (5.9)	19 (31.1)
				< 0.001
Female, n (%)	70 (47.9)	37 (43.5)	33 (54.1)	0.27
Race, Caucasian (%)	86 (58.9)	52 (61.2)	34 (55.7)	0.63
Multiple births, n (%)	22 (15.1)	14 (16.5)	8 (13.1)	0.75

^aTwo-group t-test (continuous) or Chi-square test (categorical)

infants (8 pairs of twins and 2 sets of triplets), we estimated pair-wise kinship coefficients using KING (Kinship-based INference for Genome-wide association studies) [23] for further inclusion as a covariance element to control for relatedness in subsequent mixed-effect models.

Single common variant analysis

Common variants were selected based on MAF > 0.05, missing genotype rate < 0.10, and a Hardy-Weinberg Equilibrium *p*-value > 0.01. For each common variant, we used generalized linear mixed models implemented in EMMAX (Efficient Mixed-Model Association eXpedited) to test its association with BPD, adjusting for sex,

gestational age at birth, the first two principal components, and relatedness between siblings [24]. We defined significance for this first step as a *P*-value < 0.05 (Fig. 1).

Unique variant analysis

We utilized the methods of Li, et al. [13] to identify extremely rare variants that were uniquely present in either the “affected” or “unaffected” cohorts. Briefly, we considered those variants in the cohorts that were not observed from the 1000 Genome project and dbsnp138 database, and compared them with the 1000 Genome non-synonymous variants. We used the Genomic Evolutionary Rate Profiling (GERP) score to estimate

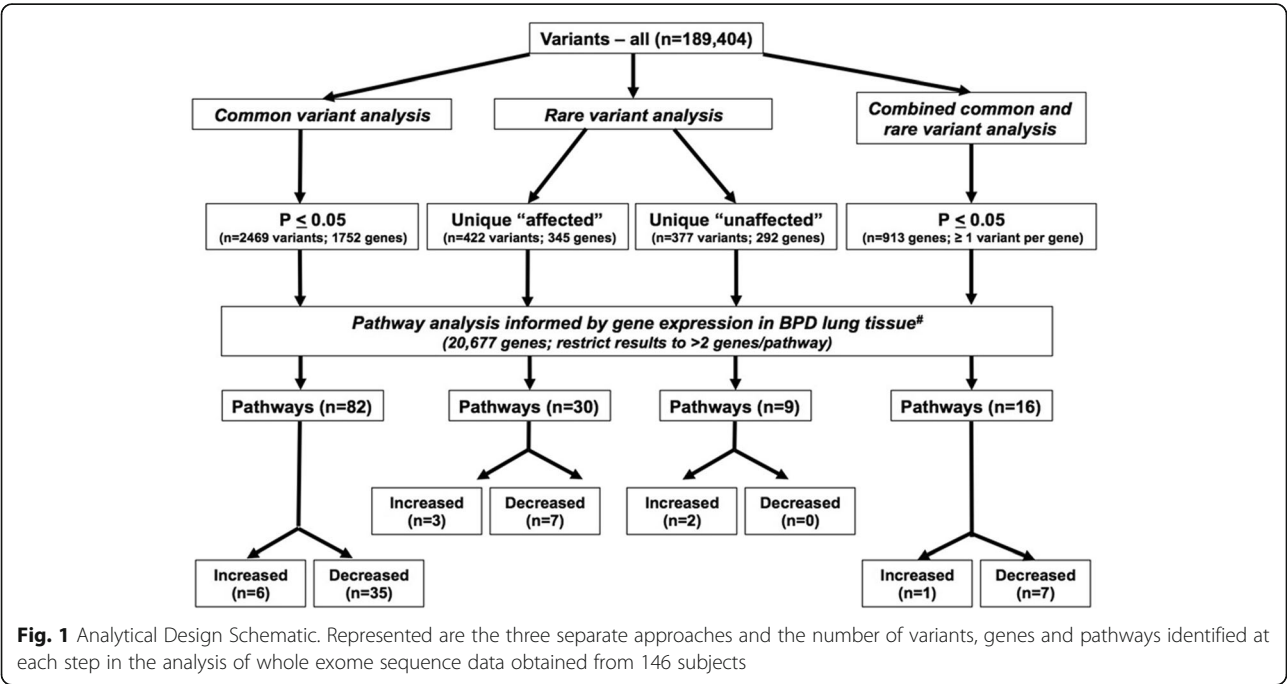


Fig. 1 Analytical Design Schematic. Represented are the three separate approaches and the number of variants, genes and pathways identified at each step in the analysis of whole exome sequence data obtained from 146 subjects

evolutionary conservation levels of these variants, where higher scores indicated greater evidence of past purifying selection [25]. To exclude stochastic mutations with low frequencies in a given population among these rare variants, we further analyzed the extremely conserved variants (GERP score > 5.83, 95% of the nonsynonymous sites in the 1000 Genome dataset) and thus considered 422 and 377 rare nonsynonymous variants identified in the 85 affected individuals and 61 unaffected individuals, respectively. When we excluded the genes that were common to both groups, we found conserved extremely rare nonsynonymous variants in 345 and 292 unique genes in the affected and unaffected cohorts, respectively (Fig. 1).

Combined rare and common variant analysis

To test the joint effect of both common and rare variants within each gene locus, we used the Fast Family-Based Sequence Kernel Association Test (FFB-SKAT) [26]. As a powerful kernel machine method, the FFB-SKAT takes linear (additive), nonlinear, and epistasis effects into account. The analysis was adjusted for sex, gestational age at birth, population stratification as well as relatedness [27, 28]. Missing genotypes were imputed using the simple mean approach [29].

Pathway analysis

Significantly overrepresented canonical pathways were identified using Ingenuity Pathway Analysis (IPA) software, using default settings (Qiagen). Canonical pathways with a $-\log(p\text{-value})$ greater than 1.301 ($p < 0.05$) were considered significant. In addition, for directional pathway analysis, we cross-referenced these results with a previously published data set defining genome-wide expression changes in postmortem lung tissues from unrelated preterm subjects with and without BPD [30]. Briefly, probe sets for the Affymetrix HG-U133 plus array were matched using Entrez Gene IDs and NetAffx (www.netaffx.com). If the Entrez Gene ID was missing, the Official Gene Symbol was used. An expression ratio comparing BPD cases to controls was included for each gene, and if no expression value was available the gene was excluded from the analysis. Canonical pathway analysis was performed and results included z-scores denoting regulatory status (activated, inhibited) of each significant pathway if 3 or more genes were represented. Complete pathway analysis results are provided in Additional file 3: Table S3, Additional file 4: Table S4, Additional file 5: Table S5 and Additional file 7: Table S7. Pathways selected for inclusion in Fig. 3 were among the most significant, were identified in more than 1 pathway analysis results or had at least one predicted directionality, and were of biological interest.

Results

Common variants

We first assessed the association between common variants, as defined as a MAF > 0.05, and BPD using EMMAX while adjusting for gestational age, sex, population stratification, and relatedness between siblings (Fig. 2a and Additional file 2: Table S2). We identified a total of 2469 common variants in 1752 genes with nominal association with BPD status, but none reaching genome-wide significance. The strongest associations identified are denoted in Additional file 2: Table S2. We then used Ingenuity Pathway Analysis to interrogate the 1752 genes with common variant association $P < 0.05$ and further informed this analysis using genome-wide expression data from lung tissue of infants with BPD [30]. Among the most significant pathways identified (Additional file 4: Table S4 and Fig. 3) included Hepatic fibrosis/hepatic stellate cell activation ($-\log p\text{-value} = 6.9$), G-protein receptor signaling ($-\log p\text{-value} = 4.1$) and caveolar-mediated endocytosis signaling ($-\log p\text{-value} = 3.4$), however BPD-specific lung tissue gene expression did not inform any directionality in these pathways (Fig. 3). Interestingly, these gene expression patterns predicted sperm motility, MAPK, protein Kinase A and corticotropin releasing hormone signaling pathways to be inhibited. Additional significant pathways identified, along with their predicted (e.g., activation/inhibition) status, are listed in Additional file 3: Table S3.

Unique variants

Using the approach described by Li et al. [13], we identified 345 genes with extremely rare variants that were uniquely present only in our severely affected subjects, 19 of which were also among the 258 identified by Li et al. in their BPD cohort (*CDH23*, *CHST10*, *DAAM1*, *DYDC1*, *FIGN*, *KIAA1468*, *MDN1*, *MYOZ1*, *NID2*, *NRXN3*, *NUP93*, *PICALM*, *ROR1*, *SETBP1*, *SYDE2*, *TBC1D1*, *TGFBI*, *YLPM1*, *ZHX3*). Further analysis of these 345 genes identified 30 significant canonical pathways of interest (Additional file 4: Table S4 and Fig. 3). Pathways representing sperm motility, protein kinase A and corticotropin releasing hormone signaling were also detected in this analysis. In addition, this analysis predicted cardiac hypertrophy signaling to be activated and gonadotropin related hormone signaling to be inhibited.

Using this approach, we also identified a total of 292 genes with extremely rare variants that were uniquely found in our unaffected subjects (Fig. 1), 9 of which were among the 182 identified by Li et al. in their non-disease cohort (*DAG1*, *FBN1*, *GABRA2*, *MRPL16*, *NCAM1*, *POLR3A*, *PPP1R3A*, *SLC38A2*, *SOX6*). Pathway analyses using these 292 genes with variants unique to this unaffected cohort identified 9 significant canonical pathways of interest (Additional file 5: Table S5 and Fig. 3). Of particular interest, the genes with extremely

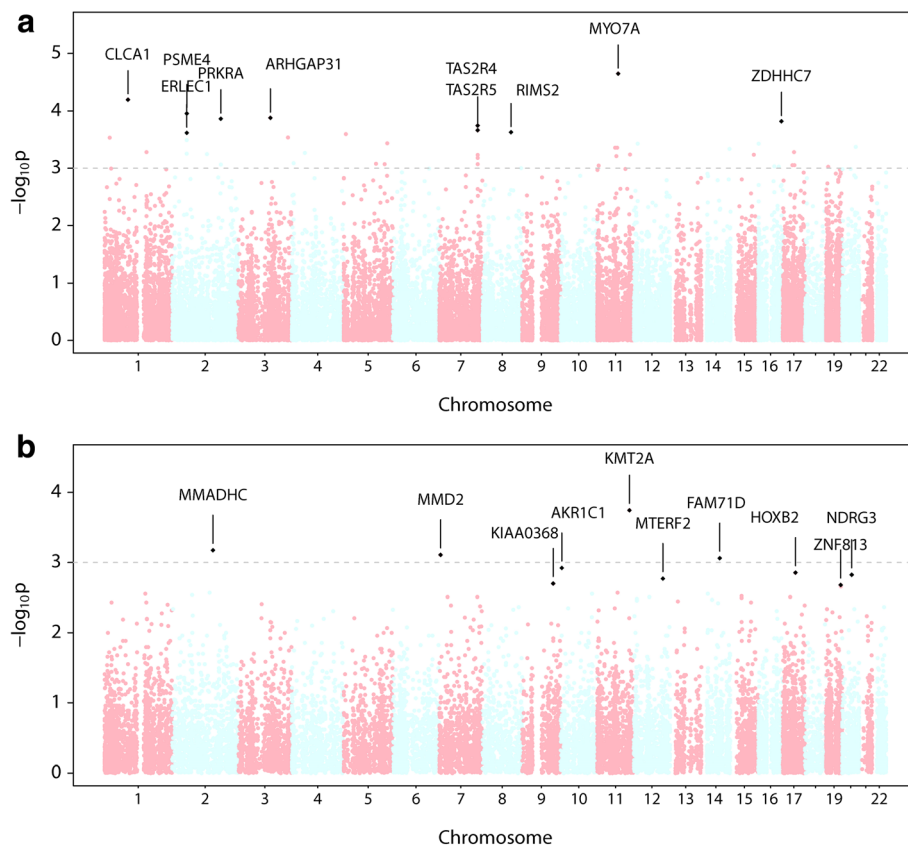


Fig. 2 Genome-wide Association Results. We performed genome-wide association analysis using either (a) all common variants (MAF > 0.05) or (b) combined rare and common variants at each individual gene locus (FFB-SKAT), comparing the affected group to the unaffected group. Manhattan plots show the strength of association for all variants, organized by chromosomal locus, with genes demonstrating greatest association indicated by name

rare variants unique to the unaffected subjects were associated with *activation* of protein kinase A signaling, whereas the genes in this pathway associated with BPD in the 2 preceding analyses predicted *inhibition* of protein kinase A signaling. These data may indicate a genetic basis for susceptibility and resilience from chronic lung disease in preterm infants.

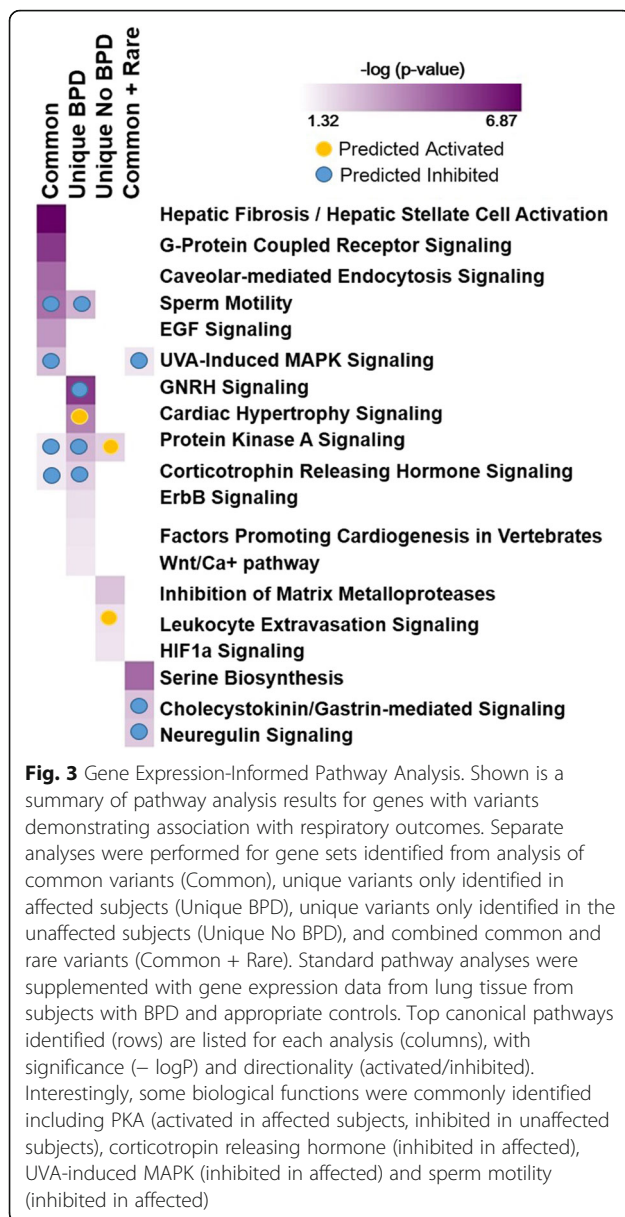
Combined rare and common variants

Hypothesizing that a combination of rare and common variants within a given gene locus are likely to contribute to risk for, or protection from, adverse respiratory outcomes, we pursued an approach to test for the combined association of all variants across individual gene loci.

Using the Fast Family-Based Sequence Kernel Association Test (FFB-SKAT) we identified 913 genes with $P \leq 0.05$ and 139 genes with $P \leq 0.01$ (Fig. 2b and Additional file 6: Table S6). Again, no genes showed evidence for significant association following adjustment for multiple testing, but the strongest associations included those for genes encoding *KMT2A* (1.8×10^{-4}),

CHRD2 (2.7×10^{-3}) and *CTSK* (2.8×10^{-3}). Interestingly, there were 16 genes identified that had unique variants in cases, and also identified by FFB-SKAT (*ANKRD31*, *ARHGAP32*, *C5orf66*, *DCLRE1A*, *FCAMR*, *KMT2A*, *NFX1*, *NRXN3*, *NSMAF*, *PIGH*, *PSAT1*, *TEX2*, *TGFB1*, *UNC45A*, *WDFY2*, *YLPM1*), and there were 12 genes identified that had unique variants in controls, and also identified by FFBSKAT (*ABCA1*, *ANKEF1*, *ANKRD17*, *FAM120A*, *MIPEP*, *MPZL2*, *PHGDH*, *SOX6*, *TPGS2*, *WDR36*, *WLS*, *YTHDC1*).

Pathway analysis using the 913 genes, informed by BPD-associated expression, identified 16 significant canonical pathways of interest (Additional file 7: Table S7). Similar to the common variant analysis, MAPK signaling was predicted to be inhibited in BPD. The single gene in this pathway with variants identified in both analyses was the epidermal growth factor receptor (*EGFR*), the principal member of the ErbB receptor tyrosine kinase family, which has been implicated in alveolar formation [31]. A related pathway, the Neuregulin signaling pathway, was among the most highly significant ($-\log p$ -value of 1.85), and was predicted to be inhibited in



BPD. Neuregulin promotes lung maturation and surfactant production, and mice deficient in neuregulin receptor (ErbB4) display BPD-like pathology [32, 33]. In this analysis, the Cholecystokinin/Gastrin-mediated Signaling pathway was also predicted to be inhibited in BPD. This is of interest given prior studies implicating gastrin-related peptide (GRP) in BPD pathogenesis [34].

For key pathways identified by our results and presented in Fig. 3, we further determined if genes with variants are expressed in human fetal lung at a time commensurate with birth at greatest risk for BPD, using data from a recently published study [35]. We found a large majority of these genes are expressed, and are therefore biologically plausible candidates for affecting

disease risk/susceptibility. Specifically, we assessed the expression of all 174 genes associated with the 19 pathways. Of 62,130 genes (coding genes plus non-coding RNAs) tested for expression, 14,500 were identified as expressed (counts per million normalized counts > 5 in control/untreated condition samples). Among the 174 genes associated with these 19 pathways, 115 genes were identified as expressed in the fetal lung data. The association between the expression of genes in this data set to greater than would be expected by chance was tested using a Chi-square test ($p < 0.001$). These data, along with an assessment of their differential expression in human BPD lung tissue [30], and their expression trajectory in early human lung development [36] are presented in Additional file 8: Table S8.

Discussion

We describe the largest exome sequencing study for respiratory outcomes of prematurity to date, and add to the body of genetic data available for this population. We focused our efforts on those subjects with extremely divergent respiratory morbidity outcomes reasoning that we would have greater power to detect highly penetrant genetic influences [16, 17, 19]. We also chose not to use conventional definitions for BPD, which are based on need for supplemental oxygen at 36 weeks post-menstrual age, since current practices for respiratory support of the premature newborn emphasize non-invasive positive pressure administration without supplemental oxygen whenever possible [1]. As would be anticipated, the cohorts were skewed with respect to gestational age, with more infants born at the earliest gestations being in the extreme affected group and those with the later gestations (though still less than 29 weeks) being more prevalent in the extreme unaffected group. Nonetheless, the common variant analysis with EMMAX and the combined common and rare variant analysis with FFB- SKAT controlled for this difference in gestational age as well as other confounders (sex and race) that could independently account for genetic differences between the two groups.

Previous candidate gene and genome-wide studies testing a million or more variants in which the MAF is $> 1\%$ have failed to identify consistent common variants, genes, or pathways that are significantly associated with respiratory morbidity following preterm birth. We report exact p -values, per published guidelines for statistical reporting, allowing readers to make their own judgments about appropriate corrections. Since WES allows for the identification and enumeration of both common and rare variants, we used several analytical approaches to test for those variants that are associated with our extreme 36-week respiratory outcomes [37].

Each of our approaches to identify disease risk-associated variants utilizes a different set of assumptions, which

probably accounts for the differing results. To gain insight into mechanistic relevance and potential function of variants with marginal associations, we integrated functional data based on gene expression from lung tissue of infants who died of BPD [30]. Results from these analyses, particularly those including rare genetic variants, identified a small set of biological pathways that warrant additional consideration (Fig. 3).

Rare variants in 10 (of 63) genes that represent the cardiac hypertrophy signaling pathway were uniquely represented in the severely affected group. This pathway was also significantly activated in BPD based upon gene expression data. Up to 25% of infants with severe BPD may develop pulmonary hypertension, which is typically attributed to elevated pulmonary vascular resistance from arrest of pulmonary vascular development and increases the potential for mortality [38–40]. However, the presence of variants in this cardiac hypertrophy pathway suggests that there may be a subset of infants whose lung disease is complicated by cardiac or pulmonary vascular dysfunction that is genetic in origin.

Rare variants in both the gonadotropin and corticotropin releasing hormone (GnRH and CRH, respectively) pathways were also uniquely identified in the affected cohort, and were inhibited in BPD based upon gene expression (Additional file 8: Table S8). Endogenous corticosteroids increase in late gestation and enhance lung maturation, and there are previously described associations between low levels of maternal unconjugated estriol in the 2nd trimester with BPD and worse neonatal outcomes, along with multiple human and animal studies that demonstrate improved lung function and outcomes with estradiol or corticosteroid administration [41–44]. In a gene expression study of peripheral blood mononuclear cells, the GnRH pathway was activated in infants who developed BPD [45]. These observations support the possible relevance of gene variants in these two pathways to the risk for more severe respiratory disease after extremely preterm birth.

Intriguingly, rare variants among separate sets of genes in the protein kinase A (PKA) signaling pathway were uniquely identified in both the affected and unaffected groups. Twelve (of 71) individual genes in the PKA pathway had variants uniquely in the affected subjects while variants in 10 different genes in this pathway were uniquely present in unaffected subjects.

Gene expression data indicate this pathway is inhibited in subjects who develop BPD, but is activated in those without BPD. While the relevance of these observations is uncertain, PKA- cAMP signaling is known to be a critical determinant in promoting respiratory epithelial cell differentiation [37, 46, 47].

Results from a more agnostic approach, combining contributions from both common and rare variants, identified different sets of genes and pathways with potential biological significance. We noticed a convergence in many of these pathways, particularly focusing upon the EGFR family, and particularly neuregulin. Variants in 7 genes of the neuregulin signaling pathway were associated with the affected group, and this pathway was predicted to be inhibited in BPD based upon gene expression. Two additional pathways implicated variation in receptors for neuregulin in the affected group: variants in 5 genes from the EGFR superfamily- related ErbB2/ErbB3 pathway and variants in 7 genes from a MAPK pathway, where *EGFR* serves as a central hub. Gene expression data again suggested inhibition of receptor signaling in both of these cases. In vitro, neuregulin has previously been reported to stimulate surfactant production mediated by erbB2 receptor phosphorylation in alveolar type II cells [32].

Furthermore, mice deficient in the neuregulin receptor *ErbB4* [33] or *EGFR* [31] display BPD- like pathology. From these data, we speculate that genetic variants that suppress neuregulin/EGFR/ErbB4 signaling may also contribute to severity of lung disease in a subset of infants with adverse respiratory outcomes. In support of this conclusion, we found variants in multiple genes of the EGF signaling pathway were present in the affected group in both the common and combined variant analysis, and in both cases gene expression predicted this pathway was inhibited in BPD.

The lack of replication in genetic studies for BPD likely stems in large part from population differences; i.e. differences in gestational ages, limited cohort sizes, population ancestry, and, perhaps most importantly, the lack of a precise definition of BPD. The platform for variant identification plays an important role, as well. Studies using a genotyping approach with common variant-focused arrays may identify genomic regions in which specific genetic information is embedded, but as most of these variants are in non-coding regions, the data typically only infer the possibility of functional variants in the region. Conversely, sequence- based genotyping, as with exon capture sequencing used in this study, overlooks functionally relevant variants in non-coding regions.

Our studies to a certain extent have similar limitations, in that we have failed to replicate many prior observed genetic associations. One of the first genome-wide screens for BPD candidate genes identified variants in *SPOCK2*, an extracellular matrix molecule, associated with multiple, independent cohorts [8]. Another relatively large genome-wide screen identified BPD-associated variants associated with *ADARB2*, *CD44* and *miR-219* [6]. We are unable to replicate these associations in our population. In fact,

the largest genome-wide screen to date failed to identify a single locus that met genome-wide significance [11], similar to ours.

However, and of potential greatest importance of our results, we have been able to successfully replicate the association of 28 genes with extremely rare variants previously associated with susceptibility or resistance to BPD following preterm birth [13]. In addition to these 28 genes, 4 pathways were implicated by unique variants identified in both studies (Additional file 9: Table S9 and Additional file 10: Table S10); three pathways were replicated in cases (Folate Transformations I, Factors Promoting Cardiogenesis in Vertebrates, and Sertoli Cell-Sertoli Cell Junction Signaling) and one in controls (Protein Kinase A Signaling). These data are consistent with the hypothesis that the number of potential genes related to lung maturity and defense systems is large, each with a relatively small effect size.

Our study has several additional limitations. Our definition of adverse respiratory outcome was based on short-term treatment before and at 36 weeks PMA. Definitions based on longer-term outcomes may help identify other risk variants underlying infants' respiratory outcomes. The size of the cohort, relative to the number of variants interrogated, limits our power to identify variants with requisite statistical significance. Additional whole genome sequencing studies with larger populations will be able to address this issue of heritability with more precision. While sequencing studies combining multiple analytic approaches may help to identify risk genes and pathological mechanisms that will, in turn, permit a more precise definition of disease phenotype, only whole genome sequencing will provide the entire range of variation and permit more in-depth look at the prevalence and relevance of non-coding region variation.

Conclusion

In summary, we used whole exome sequencing in a cohort of premature infants with the extremes of short-term respiratory outcomes, and used multiple analytical approaches to identify several genes associated with BPD. Although individual genes did not meet genome wide statistical significance, we found several candidate pathways with common themes and biological plausibility that are worthy of more in-depth exploration with additional sequencing studies.

Additional files

Additional file 1: Table S1. Sequence quality summary (DOCX 24 kb)

Additional file 2: Table S2. Top 50 Genome-wide Common SNP Association Results (DOCX 17 kb)

Additional file 3: Table S3. Significant canonical pathways represented by common variants associated with BPD. (DOCX 31 kb)

Additional file 4: Table S4. Significant canonical pathways represented by unique variants in "affected" subjects. (DOCX 28 kb)

Additional file 5: Table S5. Significant canonical pathways represented by unique variants in "unaffected" subjects. (DOCX 32 kb)

Additional file 6: Table S6. Top 50 Genes from locus-based (FFB-SKAT) Association Analysis (DOCX 16 kb)

Additional file 7: Table S7. Significant canonical pathways represented by locus-based (FFB-SKAT) association results. (DOCX 14 kb)

Additional file 8: Table S8. Gene expression in human lung tissue. (DOCX 51 kb)

Additional file 9: Table S9. Significant canonical pathways represented by unique variants in "affected" subjects from Li et al. (DOCX 14 kb)

Additional file 10: Table S10. Significant canonical pathways represented by unique variants in "unaffected" subjects from Li et al. (DOCX 22 kb)

Abbreviations

BPD: Bronchopulmonary dysplasia; EMMAX: Efficient Mixed-Model Association eXpedited; FFB-SKAT: Fast Family-Based Sequence Kernel Association Test; GERP: Genomic Evolutionary Rate Profiling; GWAS: Genome-wide association studies; IPA: Ingenuity Pathway Analysis; KING: Kinship-based INference for Genome-wide association studies; MAF: Minor allele frequency; MDS: Multidimensional scaling; NHLBI: National Heart, Lung, and Blood Institute; PMA: Postmenstrual age; PROP: Prematurity and Respiratory Outcomes Program; WES: Whole Exome Sequencing

Acknowledgements

This work was made possible by an exceptional network of highly collaborative investigators of the Prematurity and Respiratory Outcomes (PROP), who are listed in an Appendix/Supplement. In addition, the authors would like to express our gratitude to Dr. Judy Aschner, Cara Sutcliffe and Ann Walker at Vanderbilt University, and to Dr. Debbie Nickerson and Sam Phillips at the University of Washington Northwest Genomics Center, for their help in coordinating sample handling, processing, archiving and sequence data generation. We also acknowledge assistance from the PROP Data Coordinating Center at the University of Pennsylvania, including Melissa Fernando. The authors would also like to acknowledge Dr. Hugh O'Broovich at Stanford University for sharing unpublished data. Finally, we express our gratitude to the patients and families who participated in these studies.

Funding

Funding support for the PROP includes HL101456, HL101465, HL101794, HL101798, HL101800, HL101813, HL105702, and UL1TR000448. Sequencing services were provided through the RS&G Service by the Northwest Genomics Center at the University of Washington, Department of Genome Sciences, under U.S. Federal Government contract number HHSN268201100037C from the National Heart, Lung, and Blood Institute. This publication was supported by the Washington University Institute of Clinical and Translational Sciences, which is supported in part by the NIH/National Center for Advancing Translational Sciences (NCATS) CTSA grant UL1TR000448.

Availability of data and materials

Variant-level data that support the findings of this study are available, 1 year after publication, upon reasonable request by application to the corresponding authors, and following approval by the PROP Executive Committee.

Author statement

The first draft of the manuscript was written by Drs. Hamvas and Mariani. No honorarium, grant, or other form of payment was given to anyone to produce the manuscript.

Authors' contributions

AH, TJM, RF and PLB designed the study; RF, FW, SB, JM, MK, YB, AH and TJM performed data analyses; AH and TJM wrote the first draft of the manuscript; all authors read, edited and approved the manuscript.

Authors' information

PROP Investigators:

Cincinnati Children's Hospital Medical Center:

Barbara Alexander, RN, Claire Chougnet, PhD, Tari Gratton, PA, James M. Greenberg, MD, Cathy Grisby, BSN, CCRC, William Hardie, MD, Alan H. Jobe MD, PhD, Beth Koch, BHS, RRT, RPFT, Karen McDowell, MD, Kelly Thornton BS. Washington University:

Pamela Bates, CRT, RPFT, RPSGT, Claudia Cleveland, RRT, Thomas Ferkol, MD, Aaron Hamvas, MD, Julie Hoffmann, RN, Mark R. Holland, PhD, James Kemp, MD, Philip T. Levy, MD, Laura Linneman, RN, Jayne Sicard-Su, RN, Gina Simpson, RRT, CPFT, Gautam K. Singh, MD, Barbara Warner, MD.

University of California at San Francisco:

Investigators

Philip L. Ballard, MD, PhD¹, Roberta A. Ballard, MD¹, David J. Durand, MD², Eric C. Eichenwald, MD⁴, Roberta L. Keller, MD¹, Amir M. Khan, MD⁴, Leslie Lusk, MD¹, Jeffrey D. Merrill, MD³, Dennis W. Nielson, MD, PhD¹, Elizabeth E. Rogers, MD¹.

Research Staff

Jeanette M. Asselin, MS, RRT-NPS², Samantha Balan¹, Katrina Burson, RN, BSN⁴, Cheryl Chapin¹, Erna Josiah-Davis, RN, NP³, Carmen Garcia, RN, CCRP⁴, Hart Horneman¹, Rick Hinojosa, BSRT, RRT, CPFT-NPS⁴, Christopher Johnson, MBA, RRT⁴, Susan Kelley, RRT¹, Karin L. Knowles¹, M. Layne Lillie, RN, BSN⁴, Karen Martin, RN⁴, Sarah Martin, RN, BSN¹, Julie Arldt-McAlister, RN, BSN⁴, Georgia E. McDavid, RN⁴, Lori Pacello, RCP2, Shawna Rodgers, RN, BSN⁴, Daniel K. Sperry, RN⁴.¹

¹Department of Pediatrics, University of California San Francisco, San Francisco, CA; ²Children's Hospital and Research Center Oakland, Oakland, CA.

³Alta Bates Summit Medical Center, Berkeley, CA.

⁴University of Texas Health Science Center- Houston, Houston, TX.

Vanderbilt University:

Judy Aschner, MD, Amy B Beller BSN, Candice Fike, MD, Scott Guthrie, MD, Tina Hartert, MD, Nathalie Maitre, MD, Paul Moore, MD, Mark O' Hunt, Theresa J. Rogers, RN, Odessa L. Settles, RN, MSN, CM, Steven Steele, RN, Marshall Summar, MD, Sharon Wadley, BSN, RN, CLS.

University of Rochester-University at Buffalo:

Investigators

Carl D'Angio, MD, Vasanth Kumar, MD, Tom Mariani, PhD, Gloria Pryhuber, MD, Clement Ren, MD, Anne Marie Reynolds, MD, MPH, Rita M. Ryan, MD, Kristin Scheible, MD, Timothy Stevens, MD, MPH.

Technical Staff

Heidie Huyck, BS, Valerie Lunger, MS

Study Staff

Shannon Castiglione, RN, Aimee Horan, LPN, Deanna Maffet, RN, Jane O'Donnell, PNP, Michael Sacilowski, MAT, Tanya Scalise, RN, BSN, Elizabeth Werner, MPH, Jason Zayac, BS Respiratory Therapists and Nurses: Kim Bordeaux, RRT, Pam Brown, RRT, Julia Epping, AAS, RT, Lisa Flatterly-Walsh, RRT, Donna Germuga, RRT, CPFT, Nancy Jenks, RN, Mary Platt, RN, Eileen Popplewell, RRT, Sandra Prentice, CRT.

Duke University:

Kim Ciccio, RN, C. Michael Cotten, M.D., Kim Fisher, Ph.D., Jack Sharp, M.D., Judith A. Voynow, M.D.*.

*Present address, Virginia Commonwealth University.

Indiana University:

Charles Clem, RRT, Stephanie Davis, M.D., Susan Gunn, NNP, CCRC, Lauren Jewett, RN, CCRC Brenda Poindexter, M.D., M.S.#.

#Present address, University of Cincinnati.

Steering Committee Chair:

Lynn M. Taussig, MD, University of Denver.

NHLBI Program Officer:

Carol J. Blaisdell, MD.

University of Pennsylvania Data Coordinating Center:

Scarlett Bellamy, ScD, Maria Blanco, BS, Denise Cifelli, MS, Sara DeMauro, MD, Jonas Ellenberg, PhD, Rui Feng, PhD, Melissa Fernando, MPH, Howard Panitch, MD, Barbara Schmidt, MD, MSc, Pamela Shaw, PhD, Ann Tierney, BA, MS

Ethics approval and consent to participate

Studies at each of the participating PROP sites were IRB-approved and all parents consented to collection and sequencing of infant DNA and collection of clinical data.

Consent for publication

All parents consented to publication of data without identifiers.

Competing interests

All authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Pediatrics, Northwestern University, Chicago, IL, USA.

²Department of Biostatistics, University of Pennsylvania, Philadelphia, PA, USA.

³Department of Preventive Medicine, Northwestern University, Chicago, IL, USA.

⁴Department of Pediatrics, University of Rochester, Rochester, NY, USA.

⁵Center for Biomedical Informatics, Washington University, St. Louis, MO, USA.

⁶Department of Pediatrics, Duke University, Durham, NC, USA.

⁷Department of Pediatrics, University of California, San Francisco, CA, USA.

⁸Division of Neonatology and Pediatric Molecular and Personalized Medicine Program University of Rochester Medical Center, 601 Elmwood Ave, Box 850, Rochester, NY 14642, USA. ⁹Ann and Robert H. Lurie Children's Hospital of Chicago and Northwestern University, Chicago, IL, USA.

Received: 20 October 2017 Accepted: 1 October 2018

Published online: 20 October 2018

References

- Poindexter BB, Feng R, Schmidt B, Aschner JL, Ballard RA, Hamvas A, Reynolds AM, Shaw PA, Jobe AH, Prematurity and Respiratory Outcomes Program, et al. Comparisons and limitations of current definitions of bronchopulmonary dysplasia for the prematurity and respiratory outcomes program. *Ann Am Thorac Soc*. 2015;12(12):1822–30.
- Horbar JD, Carpenter JH, Badger GJ, Kenny MJ, Soll RF, Morrow KA, Buzas JS. Mortality and neonatal morbidity among infants 501 to 1500 grams from 2000 to 2009. *Pediatrics*. 2012;129(6):1019–26.
- Stoll BJ, Hansen NI, Bell EF, et al. Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993–2012. *JAMA*. 2015;314(10):1039–51.
- Bhandari V, Bizzarro MJ, Shetty A, Zhong X, Page GP, Zhang H, Ment LR, Gruen JR, Group ftnGS. Familial and genetic susceptibility to major neonatal morbidities in preterm twins. *Pediatrics*. 2006;117(6):1901–6.
- Lavoie PM, Pham C, Jang KL. Heritability of bronchopulmonary dysplasia, defined according to the consensus statement of the National Institutes of Health. *Pediatrics*. 2008;122(3):479–85.
- Ambalavanan N, Cotten CM, Page GP, Carlo WA, Murray JC, Bhattacharya S, Mariani TJ, Cuna AC, Faye-Petersen OM, Kelly D, et al. Integrated genomic analyses in bronchopulmonary dysplasia. *J Pediatr*. 2015;166(3):531–7 e513.
- Carrera P, Di Resta C, Volonteri C, Castiglioni E, Bonfiglio S, Lazarevic D, Cittaro D, Stupka E, Ferrari M, Somaschini M, et al. Exome sequencing and pathway analysis for identification of genetic variability relevant for bronchopulmonary dysplasia (BPD) in preterm newborns: a pilot study. *Clin Chim Acta*. 2015;451(Pt A):39–45.
- Hadchouel A, Durrmeyer X, Bouzigon E, Incitti R, Huusko J, Jarreau P-H, Lencle R, Demenais F, Franco-Montoya M-L, Layouni I, et al. Identification of SPOCK2 as a susceptibility gene for bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2011;184(10):1164–70.
- Lal CV, Ambalavanan N. Genetic predisposition to bronchopulmonary dysplasia. *Semin Perinatol*. 2015;39(8):584–91.
- Torgerson DG, Ballard P, Keller RL, Oh S, Huntsman S, Hu D, Eng C, Burchard E, Ballard R, TOLSURF Study Group. Ancestry and Genetic Associations with Bronchopulmonary Dysplasia in Preterm Infants. *Am J Physiol Lung Cell Mol Physiol*. 2018.
- Wang H, St. Julien KR, Stevenson DK, Hoffmann TJ, Witte JS, Lazzaroni LC, Krasnow MA, Quaintance CC, Oehlert JW, Jelliffe-Pawlowski LL, et al. A genome-wide association study (GWAS) for bronchopulmonary dysplasia. *Pediatrics*. 2013;132(2):290–7.
- Yu KH, Li J, Snyder M, Shaw GM, O'Brodovich HM. The genetic predisposition to bronchopulmonary dysplasia. *Curr Opin Pediatr*. 2016;28(3):318–23.
- Li J, Yu KH, Oehlert J, Jelliffe-Pawlowski LL, Gould JB, Stevenson DK, Snyder M, Shaw GM, O'Brodovich HM. Exome sequencing of neonatal blood spots and the identification of genes implicated in bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2015;192(5):589–96.

14. Maitre NL, Ballard RA, Ellenberg JH, Davis SD, Greenberg JM, Hamvas A, Pryhuber GS. Prematurity, respiratory outcomes P Respiratory consequences of prematurity: evolution of a diagnosis and development of a comprehensive approach. *J Perinatol* 2015; 35(5):313–321.
15. Pryhuber GS, Maitre NL, Ballard RA, Cifelli D, Davis SD, Ellenberg JH, Greenberg JM, Kemp J, Mariani TJ, Panitch H, et al. Prematurity and respiratory outcomes program (PROP): study protocol of a prospective multicenter study of respiratory outcomes of preterm infants in the United States. *BMC Pediatr*. 2015;15:37.
16. Bruse S, Moreau M, Bromberg Y, Jang JH, Wang N, Ha H, Picchi M, Lin Y, Langley RJ, Qualls C, et al. Whole exome sequencing identifies novel candidate genes that modify chronic obstructive pulmonary disease susceptibility. *Hum Genomics*. 2016;10:1.
17. Emond MJ, Louie T, Emerson J, Chong JX, Mathias RA, Knowles MR, Rieder MJ, Tabor HK, Nickerson DA, Barnes KC, et al. Exome sequencing of phenotypic extremes identifies CAV2 and TMC6 as interacting modifiers of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *PLoS Genet*. 2015;11(6):e1005273.
18. Sardell RJ, Bailey JN, Courtenay MD, Whitehead P, Laux RA, Adams LD, Fortun JA, Brantley MA Jr, Kovach JL, Schwartz SG, et al. Whole exome sequencing of extreme age-related macular degeneration phenotypes. *Mol Vis*. 2016;22:1062–76.
19. Shtir C, Aldahmesh MA, Al-Dahmesh S, Abboud E, Alkuraya H, Abouammoh MA, Nowailaty SR, Al-Thubaiti G, Naim EA, ALYounes B, et al. Exome-based case-control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. *Hum Genet*. 2016;135(2):193–200.
20. Uzun A, Schuster J, McGonnigal B, Schorl C, Dewan A, Padbury J. Targeted sequencing and meta-analysis of preterm birth. *PLoS One*. 2016;11(5):e0155021.
21. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010; 38(16):e164.
22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–75.
23. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867–73.
24. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44(7):821–4.
25. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 2010;6(12):e1001025.
26. Svishcheva GR, Belongova NM, Axenovich TL. FBSKAT: fast family-based sequence kernel association test. *PLoS One*. 2014;9(6):e99407.
27. Schaid DJ, McDonnell SK, Sinnwell JP, Thibodeau SN. Multiple genetic variant association testing by collapsing and kernel methods with pedigree or population structured data. *Genet Epidemiol*. 2013;37(5):409–18.
28. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*. 2011;89(1):82–93.
29. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. *Am J Hum Genet*. 2013;92(6):841–53.
30. Bhattacharya S, Go D, Krenitsky DL, Huyck HL, Solleti SK, Lunger VA, Metlay L, Srisuma S, Wert SE, Mariani TJ, et al. Genome-wide transcriptional profiling reveals connective tissue mast cell accumulation in bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2012;186(4):349–58.
31. Li J, Li Y, He H, Liu C, Li W, Xie L, Zhang Y. Csk/Src/EGFR signaling regulates migration of myofibroblasts and alveolarization. *Am J Physiol Lung Cell Mol Physiol*. 2016;310(6):L562–71.
32. Dammann CE, Nielsen HC, Carraway KL 3rd. Role of neuregulin-1 beta in the developing lung. *Am J Respir Crit Care Med*. 2003;167(12):1711–6.
33. Purevdorj E, Zscheppang K, Hoymann HG, Braun A, von Mayersbach D, Brinkhaus MJ, Schmiedl A, Dammann CE. ErbB4 deletion leads to changes in lung function and structure similar to bronchopulmonary dysplasia. *Am J Physiol Lung Cell Mol Physiol* 2008; 294(3):L516–L522.
34. Degan S, Lopez GY, Kevill K, Sunday ME. Gastrin-releasing peptide, immune responses, and lung disease. *Ann N Y Acad Sci*. 2008;1144:136–47.
35. Barrette AM, Roberts JK, Chapin C, Egan EA, Segal MR, Oses-Prieto JA, Chand S, Burlingame AL, Ballard PL. Antiinflammatory effects of budesonide in human fetal lung. *Am J Respir Cell Mol Biol*. 2016;55(5):623–32.
36. Kho AT, Bhattacharya S, Tantisira KG, Carey VJ, Gaedigk R, Leeder JS, Kohane IS, Weiss ST, Mariani TJ. Transcriptomic analysis of human lung development. *Am J Respir Crit Care Med*. 2010;181(1):54–63.
37. Gonzales LW, Guttentag SH, Wade KC, Postle AD, Ballard PL. Differentiation of human pulmonary type II cells in vitro by glucocorticoid plus cAMP. *Am J Physiol Lung Cell Mol Physiol*. 2002;283(5):L940–51.
38. Hilgendorff A, Apitz C, Bonnet D, Hansmann G. Pulmonary hypertension associated with acute or chronic lung diseases in the preterm and term neonate and infant. The European Paediatric Pulmonary Vascular Disease Network, endorsed by ISHLT and DGPK. *Heart*. 2016;102(Suppl 2):ii49–56.
39. Khemani E, McElhinney DB, Rhein L, Andrade O, Lacro RV, Thomas KC, Mullen MP. Pulmonary artery hypertension in formerly premature infants with bronchopulmonary dysplasia: clinical features and outcomes in the surfactant era. *Pediatrics*. 2007;120(6):1260–9.
40. Kim DH, Kim HS, Choi CW, Kim EK, Kim BI, Choi JH. Risk factors for pulmonary artery hypertension in preterm infants with moderate or severe bronchopulmonary dysplasia. *Neonatology*. 2012;101(1):40–6.
41. Doyle LW, Halliday HL, Ehrenkranz RA, Davis PG, Sinclair JC. An update on the impact of postnatal systemic corticosteroids on mortality and cerebral palsy in preterm infants: effect modification by risk of bronchopulmonary dysplasia. *J Pediatr*. 2014;165(6):1258–60.
42. Jelliffe-Pawlowski LL, Shaw GM, Stevenson DK, Oehlert JW, Quaintance C, Santos AJ, Baer RJ, Currier RJ, O'Brodovich HM, Gould JB. Risk of bronchopulmonary dysplasia by second trimester maternal serum levels of alpha-fetoprotein, human chorionic gonadotropin, and unconjugated Estradiol. *Pediatr Res*. 2012;71(4 0 1):399–406.
43. McCurnin DC, Pierce RA, Willis BC, Chang LY, Yoder BA, Yuhanna IS, Ballard PL, Clyman RI, Waleh N, Maniscalco W, et al. Postnatal estradiol up-regulates lung nitric oxide synthases and improves lung function in bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2009;179(6):492–500.
44. Nykänen P, Anttila E, Heinonen K, Hallman M, Voutilainen R. Early hypoadrenalism in premature infants at risk for bronchopulmonary dysplasia or death. *Acta Paediatr*. 2007;96(11):1600–5.
45. Pietrzyk JJ, Kwint P, Wollen EJ, Bik-Multanowski M, Madetko-Talowska A, Günther C-C, Jagla M, Tomasik T, Saugstad OD. Gene expression profiling in preterm infants: new aspects of bronchopulmonary dysplasia development. *PLoS One*. 2013;8(10):e78585.
46. Ballard PL, Gonzales LW, Williams MC, Roberts JM, Jacobs MM. Differentiation of type II cells during explant culture of human fetal lung is accelerated by endogenous prostanoids and adenosine 3',5'-monophosphate. *Endocrinology*. 1991;128(6):2916–24.
47. Wang Y, Maciejewski BS, Lee N, Silbert O, McKnight NL, Frangos JA, Sanchez-Esteban J. Strain-induced fetal type II epithelial cell differentiation is mediated via cAMP-PKA-dependent signaling pathway. *Am J Physiol Lung Cell Mol Physiol*. 2006;291(4):L820–7.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

