

Expression network analysis reveals cord blood vitamin D-associated genes affecting risk of early life wheeze

Hooman Mirzakhani,¹ Amal A Al-Garawi,¹ Vincent J Carey,¹ Weiliang Qiu,¹ Augusto A Litonjua,² Scott T Weiss¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/thoraxjnl-2018-211962>).

¹Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA
²Division of Pediatric Pulmonary Medicine, Department of Pediatrics, Golisano Children's Hospital at Strong, University of Rochester Medical Center, Rochester, New York, USA

Correspondence to

Dr Scott T Weiss, Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA; restw@channing.harvard.edu

HM and AAA-G contributed equally.

Received 19 April 2018
Revised 19 June 2018
Accepted 25 June 2018
Published Online First 18 July 2018



© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Mirzakhani H, Al-Garawi AA, Carey VJ, *et al.* *Thorax* 2019;**74**:200–202.

ABSTRACT

Cord blood 25-hydroxyvitamin D (25OHD) has been reported in association with risk of early life recurrent wheeze. In a subset of infants who participated in the Vitamin D Antenatal Asthma Reduction Trial, we demonstrated that higher cord blood 25OHD at birth (>31 ng/mL) was associated with a reduced risk of recurrent wheeze in the first year of life. We then identified a module of co-expressed genes associated with cord blood 25OHD levels >31 ng/mL. Genes in this module are involved in biological and immune pathways related to development and progression of asthma pathogenesis including the Notch1 and transforming growth factor-beta signalling pathways.

INTRODUCTION

Recurrent wheeze in infancy is associated with increased risk of asthma later in childhood.¹ Several studies have reported an inverse relationship between prenatal vitamin D (25OHD) levels and early life wheezing.^{2–3} The potential biological mechanisms underpinning these associations are unknown and we sought to investigate them in a subset of neonates born to participants in the Vitamin D Antenatal Asthma Reduction Trial (VDAART). VDAART is a randomised, double blind, placebo-controlled clinical trial of vitamin D supplementation (4000 International Units (IU) plus a multivitamin with 400 IU daily) versus placebo (placebo pill plus a multivitamin with 400 IU daily) in pregnant women to prevent asthma and recurrent wheeze in their children (online supplementary file 1).

METHODS

Study design

A total of 56 cases with recurrent wheezing (≥2 episodes) were matched with 113 healthy infants with no wheeze in the first year of life. The matching procedure employed information on study centre, child's race and maternal asthma to form optimally matched strata.⁴ Cord blood vitamin D (CBVD) levels were measured in umbilical cord blood (CB) plasma at birth using liquid chromatography-tandem mass spectrometry (LC-MS/MS). Total RNA was isolated from umbilical CB using the Paxgene Blood RNA Kit (Qiagen) and qualified per manufacturer's protocol. CB gene expression data were obtained using the Affymetrix GeneChip Human 1.0ST and normalised. Batch-corrected gene expressions were used in differential

expression analysis and restricted to probes derived from autosomal chromosomes. Probes with overall between-sample variance in the first quintile of the distribution of the variances across probes were excluded. This approach resulted in a final expression set of 15 252 probes from 169 samples.

Statistical analysis

Cord blood vitamin D levels and wheezing status in first year of life

Conditional logistic regression, incorporating matched strata, was used to examine the association between CBVD levels and wheezing status in the first year of life. CBVD levels were categorised into quartiles with the highest quartile representing sufficiency,⁵ that is, <13, 13–22, 22–31 and >31 ng/mL.

Gene expression, gene co-expression and TF network analysis

In linear models adjusting for child race, gestational age at delivery and gender, genes with differential expressions were identified (nominal p value from the moderated t-statistics <0.05). These CBVD-associated genes were further clustered into modules of highly interconnected genes using weighted gene co-expression network analysis (WGCNA).⁶ The first principal component ('eigengene') for the module expression matrix was computed and used in conditional logistic regression to measure the association with infant wheezing status adjusting for quartiles of CBVD level. Any module exhibiting significant correlation was chosen for further characterisation. The MetaCore platform was further used to identify transcriptional factor (TF) subnetworks of the module and key hubs. TF subnetworks were ranked by a false discovery rate (FDR) <0.05. Enrichment Analyses using InnateDB and Consolidated Pathways from GeneMANIA were conducted to determine the Gene Ontology functional terms of module TFs (FDR<0.05). The online supplementary file 1 provides detailed information on the study design.

RESULTS

Study population's characteristics

Infants had a median gestational age of 39 weeks at delivery and consisted of 83 girls and 86 boys. A total of 102 infants were of African American background, while 67 infants were white or of other races (online supplementary file 2: online supplementary table S1). After the matching process, we



Table 1 The association of cord blood vitamin D levels and the identified blue module with risk of wheezing status in 1st year of life

Predictor	Coefficient	OR (95% CI)	P value
Vitamin D level (ng/mL)			
<13	Reference		
13–22	–0.82	0.44 (0.17 to 1.10)	0.08
22–31	–0.54	0.58 (0.22 to 1.50)	0.26
>31	–1.08	0.34 (0.12 to 0.96)	0.04*
Blue module (EG)	–0.42	0.65 (0.46 to 0.94)	0.02*

*P<0.05.

EG, eigenvalue.

examined the distribution of subjects across trial arms and found an even distribution (84 and 85 children from intervention and placebo arms, respectively). There was also no difference in marginal distributions by case and control status ($p=0.16$).

Cord blood vitamin D levels and wheezing status in first year of life

A protective association between higher levels of CBVD and risk of recurrent wheeze was observed (p for trend across increasing CBVD quartiles=0.031). The top-to-bottom quartile comparison (CBVD >31 ng/mL vs CBVD <13 ng/mL) yielded an estimated OR for recurrent wheeze of 0.30 (95%CI 0.11 to 0.88). Trial arm was not a predictor of the early life recurrent wheeze ($p=0.12$).

Gene expression, gene co-expression and TF network analysis

Linear regression analysis identified 1327 genes with differential expression (796 down-regulated) potentially associated with CBVD quartiles (nominal $p<0.05$, online supplementary file 3). WGCNA clustered these genes into two co-expression modules named ‘blue’ and ‘turquoise’ that included 73 and 231 co-expressed genes, respectively. A ‘grey’ module was also generated, representing those genes that did not fit into either of the other two modules (1023 transcripts). Conditional logistic regression further confirmed the association between only the blue module eigengene and lower risk of wheezing in the first year (aOR=0.65, 95%CI 0.46 to 0.94, $p=0.02$; table 1), adjusting for quartiles of CBVD levels.

The blue module was enriched for several biological pathways contributing to the development and progression of asthma, including thymic stromal lymphopoietin (TSLP), T-cell receptor (TCR) and B-cell receptor (BCR) signalling as well as immune initiating pathways, such as interleukin (IL)-3, IL-4, IL-5 and granulocyte macrophage colony-stimulating factor (GM-CSF) (all FDR<0.05, table 2 and online supplementary file 2: online supplementary figure S1). TF network analysis identified 30 subnetworks of TFs that acted on the blue module genes (all FDR<0.01, online supplementary file 2: online supplementary table S2). One sub-network containing seven total nodes with six seed nodes from the blue module included the vitamin D receptor (VDR), representing interactions between the VDR TF and the blue module including the transforming growth factor (TGF)-beta signalling pathway (online supplementary file 2: online supplementary figure S2).

Sixty-three out of 73 genes in the blue module (86.3%) made up the largest connected component (LCC), of which, 13 genes (18%) including *Lyn* and *Notch1* have previously reported in association with asthma (online supplementary file 4).^{7,8} *Lyn* demonstrated the highest degree of connectivity in the LCC

Table 2 Blue module pathway enrichment analysis using Innate DB (top 15 enriched pathways)

Pathway name	FDR	Gene symbols
IL-4 signalling	0.01	JAK1; PTPN6
LIF signalling	<0.0001	CREBBP; JAK1
TSLP		DBNL; DOCK5; JAK1; KDM3B; LYN; PTPN6; STK10; WIPF1; ZZEF1
BCR	<0.0001	ATP2B4; CHST15; DBNL; LCP2; LYN; PTPN6; RAPGEF1; RPS6KA1; SH3BP2
TCR	<0.0001	CREBBP; DBNL; DIAPH1; ELMO1; LCP2; LYN; PTPN6; RAPGEF1; SH2D3C; SH3BP2; WIPF1
IL-5	<0.001	JAK1; LYN; PTPN6; RAPGEF1; RPS6KA1
IL-4	0.01	CREBBP; JAK1; PTPN6; SPI1
EGFR1	0.01	ACTN4; EPS15L1; IQGAP1; JAK1; LYN; PTPN6; PTPRE; RPS6KA1; SH2D3C
IL-3	0.01	JAK1; LYN; PTPN6; RAPGEF1
RANKL	0.01	JAK1; LYN; PTPN6; SPI1
Kit receptor	0.016	LYN; PTPN6; RPS6KA1; WIPF1
Prolactin	0.016	CREBBP; JAK1; LYN; PTPN6
IL-2	0.029	JAK1; LYN; PTPN6
IL-6	0.03	CREBBP; JAK1; LYN
IL-11	0.04	JAK1; RPS6KA1

BCR, B-cell receptor; EGFR, epidermal growth factor receptor; FDR, false discovery rate; IL, interleukin; LIF, leukaemia inhibitory factor; RANKL, receptor activator of nuclear factor kappa-B ligand; TCR, T-cell receptor; TSLP, thymic stromal lymphopoietin.

with 27 interactions. Shortest path network analysis demonstrated that *Notch1* was connected to both module subsets of prior known asthma genes and putative novel genes via *Lyn* and *PHF12* (figure 1, online supplementary file 2: online supplementary table S3).

DISCUSSION

In this study, higher CBVD levels (>31 ng/mL) were associated with a 70% lower risk of ‘recurrent wheeze’ in the first year of life. A module of 73 co-expressed genes was associated with both CBVD levels and wheezing status in the first year of life and a 35% reduction in the risk of ‘recurrent wheeze’ in the first year of life.

Our results are in line with Rochat and colleagues who demonstrated that vitamin D intake increased targeted CB gene expressions.⁹ The co-expressed clustered gene among CBVD-associated gene signatures in this study was largely characterised by immune-related genes. The observed VDR pathway connectivity within the module provides further evidence on a plausible preventive role of vitamin D. Among the 63 interconnected module genes, *Lyn* was a regulatory hub and enriched in several immune pathways involved in cytokines regulation with a known contribution to asthma pathogenesis (table 2). *Lyn* is an inhibitor of Lyn kinase and expressed in haemopoietic cells. Previous studies suggest an essential role of *Lyn* in IL5R alpha signalling and the antiapoptotic effect of IL-5 in eosinophils.¹⁰

Known asthma genes were connected with putative novel genes in the blue module through *Notch1*. Several studies demonstrated the role of Notch signalling in asthma pathogenesis and cell fate and cell communication.⁸ Among the putative blue module novel genes, *PHF12* was another hub gene. Mutational alteration of PHD finger proteins has been related to immune disease.¹¹ The precise role and interaction of *Lyn*,

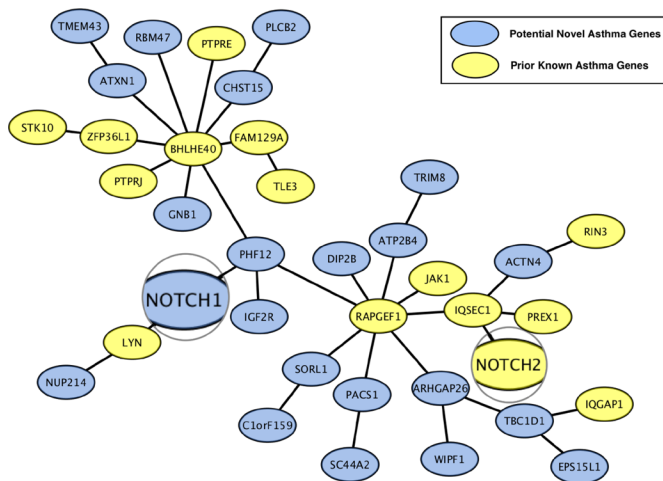


Figure 1 The linkage of prior known asthma genes in the blue module to potential novel genes using the shortest path connectivity algorithm. NOTCH signalling pathway (*Notch1*) along with potential novel *PHF12* gene plays the key role in connecting asthma-related genes to novels ones.

Notch1 and PHF12 signalling pathways (figure 1) in promoting early life wheezing and pathogenesis of asthma later in life should be investigated.

This study used a novel systematic approach to identify a co-expression network associated with first year of life wheezing. The gene set used for network analysis consisted of genes whose expression was nominally significantly associated with CBVD concentrations. This list of genes likely includes genes whose associations are artifactual. However, identification of co-expression modules, assessment of module association with disease risk and TF network analysis were conducted with multiplicity corrections. This systems biology approach to identify TF signalling networks that might play important roles in early life recurrent wheezing should be followed up with independent confirmation and functional tests of the mechanistic hypotheses implied by these findings.

VDAART registration

This study is a secondary analysis on the data from Vitamin D Antenatal Asthma Reduction Trial (VDAART). VDAART

is registered on clinicaltrials.gov with identification number NCT00920621. The trial records and results are posted.

Contributors STW: conceived the study and was the principle investigator. AAL: assisted in the design of the study and VDAART. HM, AAA-G, VJC and WQ: performed the analysis of the data. HM, VJC and AAA-G: contributed to the discussion of the statistical analysis plan and interpretation of study results. HM and AAA-G: wrote the manuscript with all other authors contributed to the relevant sections and approved the final manuscript. HM and STW: revised the manuscript.

Funding VDAART was supported by U01HL091528 from the National Heart, Lung, and Blood Institute now R01HL091528. AAG and HM are supported by T32-HL00742707. HM was also supported by L30-HL129467-01. VJC supported in part by U41 HG004059 from the National Human Genome Research Institute.

Disclaimer The corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.

Competing interests None declared.

Patient consent Obtained.

Ethics approval VDAART IRB approval was obtained from each of the three clinical centers and the Data Coordinating Center, that is, Washington University in St. Louis, Kaiser Health Care San Diego, Boston Medical Center and Brigham and Women’s Hospital in Boston.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Stern DA, Morgan WJ, Halonen M, et al. Wheezing and bronchial hyper-responsiveness in early childhood as predictors of newly diagnosed asthma in early adulthood: a longitudinal birth-cohort study. *Lancet* 2008;372:1058–64 Inpress.
- 2 Baiz N, Dargent-Molina P, Wark JD, et al. Cord serum 25-hydroxyvitamin D and risk of early childhood transient wheezing and atopic dermatitis. *J Allergy Clin Immunol* 2014;133:147–53.
- 3 Camargo CA, Ingham T, Wickens K, et al. Cord-blood 25-hydroxyvitamin D levels and risk of respiratory infection, wheezing, and asthma. *Pediatrics* 2011;127:e180–7.
- 4 Hansen BB, Klopfer SO. Optimal full matching and related designs via network flows. *JCGS* 2006;15:609–27.
- 5 Holick MF, Binkley NC, Bischoff-Ferrari HA, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011;96:1911–30.
- 6 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.
- 7 Ingley E. Functions of the Lyn tyrosine kinase in health and disease. *Cell Commun Signal* 2012;10:21.
- 8 Xu K, Moghal N, Egan SE. Notch signaling in lung development and disease. *Adv Exp Med Biol* 2012;727:89–98.
- 9 Rochat MK, Ege MJ, Plabst D, et al. Maternal vitamin D intake during pregnancy increases gene expression of ILT3 and ILT4 in cord blood. *Clin Exp Allergy* 2010;40:786–94.
- 10 Kouro T, Takatsu K. IL-5- and eosinophil-mediated inflammation: from discovery to therapy. *Int Immunol* 2009;21:1303–9.
- 11 Musselman CA, Kutateladze TG. PHD fingers: epigenetic effectors and potential drug targets. *Mol Interv* 2009;9:314–23.