

# Vitamin D prenatal programming of childhood metabolomics profiles at age 3 y

Kevin Blighe,<sup>1</sup> Bo L Chawes,<sup>1,2</sup> Rachel S Kelly,<sup>1</sup> Hooman Mirzakhani,<sup>1</sup> Michael McGeachie,<sup>1</sup> Augusto A Litonjua,<sup>1</sup> Scott T Weiss,<sup>1</sup> and Jessica A Lasky-Su<sup>1</sup>

<sup>1</sup>Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; and <sup>2</sup>Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark

## ABSTRACT

**Background:** Vitamin D deficiency is implicated in a range of common complex diseases that may be prevented by gestational vitamin D repletion. Understanding the metabolic mechanisms related to in utero vitamin D exposure may therefore shed light on complex disease susceptibility.

**Objective:** The goal was to analyze the programming role of in utero vitamin D exposure on children's metabolomics profiles.

**Design:** First, unsupervised clustering was done with plasma metabolomics profiles from a case-control subset of 245 children aged 3 y with and without asthma from the Vitamin D Antenatal Asthma Reduction Trial (VDAART), in which pregnant women were randomly assigned to vitamin D supplementation or placebo. Thereafter, we analyzed the influence of maternal pre- and post-supplement vitamin D concentrations on cluster membership. Finally, we used the metabolites driving the clustering of children to identify the dominant metabolic pathways that were influential in each cluster.

**Results:** We identified 3 clusters of children characterized by 1) high concentrations of fatty acids and amines and low maternal postsupplement vitamin D (mean  $\pm$  SD;  $27.5 \pm 11.0$  ng/mL), 2) high concentrations of amines, moderate concentrations of fatty acids, and normal maternal postsupplement vitamin D ( $34.0 \pm 14.1$  ng/mL), and 3) low concentrations of fatty acids, amines, and normal maternal postsupplement vitamin D ( $35.2 \pm 15.9$  ng/mL). Adjusting for sample storage time, maternal age and education, and both child asthma and vitamin D concentration at age 3 y did not modify the association between maternal postsupplement vitamin D and cluster membership ( $P = 0.0014$ ). Maternal pre-supplement vitamin D did not influence cluster membership, whereas the combination of pre- and postsupplement concentrations did ( $P = 0.03$ ).

**Conclusions:** Young children can be clustered into distinct biologically meaningful groups by their metabolomics profiles. The clusters differed in concentrations of inflammatory mediators, and cluster membership was influenced by in utero vitamin D exposure, suggesting a prenatal programming role of vitamin D on the child's metabolome. This trial was registered at clinicaltrials.gov as NCT00920621.

*Am J Clin Nutr* 2017;106:1092–9.

**Keywords:** airway inflammation, metabolomics, polyunsaturated fatty acids, prenatal programming, vitamin D

## INTRODUCTION

Vitamin D deficiency has now been linked to numerous complex diseases (1, 2), including respiratory diseases (3), cardiovascular diseases (4), cancers (5), depression (6), rheumatoid arthritis (7, 8), autoimmunity (9), and other inflammatory processes (10). An estimated 50% of the world's population is deficient in vitamin D (11), and there is therefore growing interest in this micronutrient and the potential for its use in health intervention and disease prevention (2, 12).

Asthma is the one of the earliest occurring chronic complex diseases (13) for which observational and mechanistic evidence suggest a role of vitamin D in its pathogenesis (14). We therefore recently conducted 2 independent randomized clinical trials of gestational vitamin D supplementation and meta-analyzed the trial results, showing that vitamin D compared with placebo decreased offspring's asthma risk at age 3 y by  $\geq 25\%$  (HM Wolsk, AA Litonjua, BL Chawes, BW Hollis, J Waage, K Bønnelykke, H Bisgaard, ST Weiss, unpublished results, 2017). These results suggest a prenatal programming role of vitamin D on the development of childhood asthma, which may be captured in the child's metabolism. However, to our knowledge, no previous study has examined the effect of in utero vitamin D exposure on the offspring's metabolomics profile.

Most of the previous vitamin D studies have only analyzed blood concentrations of the compound itself (1) in relation to inflammatory processes (15, 16) but not the large amount of

---

Supported by NIH grant HL-091528 to the Vitamin D Antenatal Asthma Reduction Trial; by National Heart, Lung, and Blood Institute R01 grant "Integrative metabolomics of asthma severity" 1R01-HL-123915-01 to JAL-S; and Electronic Medical Records and Genomics network grant U01-HG-008685 to KB.

Supplemental Figures 1–4 and Supplemental Table 1 are available from the "Online Supporting Material" link in the online posting of the article and from the same link in the online table of contents at <http://ajcn.nutrition.org>.

KB and BLC are joint first authors.

Address correspondence to JAL-S (e-mail: [rejas@channing.harvard.edu](mailto:rejas@channing.harvard.edu)).

Abbreviations used: DAG, directed acyclic graph; UPLC-MS/MS, ultra-performance liquid chromatography-tandem mass spectroscopy; VDAART, Vitamin D Antenatal Asthma Reduction Trial.

Received April 6, 2017. Accepted for publication July 18, 2017.

First published online August 23, 2017; doi: <https://doi.org/10.3945/ajcn.117.158220>.

downstream metabolites that are influenced by or interact with vitamin D. Metabolomics is a promising approach for addressing this gap in knowledge, because it is a high-throughput technology for measuring and characterizing all small-molecule metabolites in a biological sample, which represent the downstream products of a subject's genetic makeup and environmental exposures such as vitamin D (17).

We therefore aimed to evaluate the possible prenatal programming role of vitamin D by leveraging plasma metabolomic profiles from a nested case-control subset of 3-y-old children with and without asthma from the Vitamin D Antenatal Asthma Reduction Trial (VDAART; [clinicaltrials.gov](http://clinicaltrials.gov) NCT00920621) (18). The trial was approved by the Institutional Review Boards of the participating institutions and at Brigham and Women's Hospital. We used a network-based approach to optimally cluster children by their metabolomic profile at age 3 y and explored the influence of in utero vitamin D exposure on the formation of these clusters of children.

## METHODS

### The VDAART clinical trial: study participants

VDAART recruited nonsmoking pregnant women between 10 and 18 wk of gestation who reported a history of asthma, eczema, or allergic rhinitis, or who had conceived the child with a man with a history of such diseases (19). These women were randomly assigned 1:1 to a daily dose of 4000 IU vitamin D<sub>3</sub> (cholecalciferol) plus a multivitamin containing 400 IU cholecalciferol, or a matching placebo tablet plus a multivitamin containing 400 IU cholecalciferol (i.e., 4400 IU or 400 IU cholecalciferol supplement/d until delivery). The women were monitored monthly during pregnancy, and follow-up of the children occurred by telephone interviews every 3 mo and by in-person visits at ages 1, 2, and 3 y. The primary outcome of the vitamin D randomized controlled trial was asthma or recurrent wheeze at age 3 y.

The current study investigated a nested case-control population within VDAART of children with and without asthma at age 3 y diagnosed as described previously (19). The selection was done by using 1:2 frequency matching of children with and without asthma with age, sex, race, and center as the matching factors. Refer to **Supplemental Figure 1** for a participant flowchart that shows how this nested population was selected.

### Vitamin D assessments

Blood was drawn from the mothers at trial entry before supplementation (10–18 wk of gestation) and at the third trimester visit after supplementation (32–38 wk of gestation) for measurement of circulating plasma concentrations of 25-hydroxyvitamin D, determined by the DiaSorin LIAISON chemiluminescence immunoassay (20). The same method was used to measure plasma 25-hydroxyvitamin D concentrations were measured in the children at age 3 y.

### Plasma metabolomics profiling

Blood samples taken from the selected case-control subset of children at age 3 y were used for metabolomic profiling. Sample preparation was conducted according to previously described methods (17). Nontargeted global metabolomic profiles were

generated at Metabolon Inc. by using ultra-performance liquid chromatography–tandem mass spectroscopy (UPLC-MS/MS). The following 4 platforms were used to detect a comprehensive list of metabolites throughout the metabolome: 1) UPLC-MS/MS under positive ionization, 2) UPLC-MS/MS under negative ionization, 3) UPLC-MS/MS, polar platform (negative ionization), and 4) gas chromatography-MS. Metabolites were identified by their *m/z*, retention time, and through a comparison to library entries of purified known standards. In brief, the data preparation steps and quality control procedures were as follows: removing metabolites if amount in the quality control samples had CV >25%, missingness >10% across test samples, or no variability across test samples based on the IQR; removing samples with metabolite missingness >10%; and filtering out unidentified and unknown metabolites and those classified as xenobiotic chemicals.

### Use of metabolomic profiles to identify distinct clusters of children

Children were segregated into clusters according to the similarity of their metabolomic profile. The optimal number of clusters of children was determined with a 250× bootstrapped clustering function that uses partitioning around medoids to identify optimal cluster segregation. This bootstrapped function cycles through the number of potential cluster solutions ranging from  $k = 1$  to 25, calculating, at each  $k$ , the “gap” between the goodness of fit of the null distribution and that of the supplied data matrix, along with the SE and within-cluster sum of squares. In determining cluster segregation for each cluster solution, the algorithm picks a vector of raw metabolite concentrations as coordinates whose mean dissimilarity to all other coordinates in the cluster is minimal based on Euclidean distance. These data points, which are the most centrally located points in each cluster, are called the medoids. The gap statistic (21) indicated an ideal cluster range of 2 to 5, with 2 having a relatively large SE in this cluster range. We eventually decided that a 3-cluster solution was most informative based on having a low SE and using further information from, primarily, the silhouette coefficient (22) and also the elbow method (see **Supplemental Figure 2**).

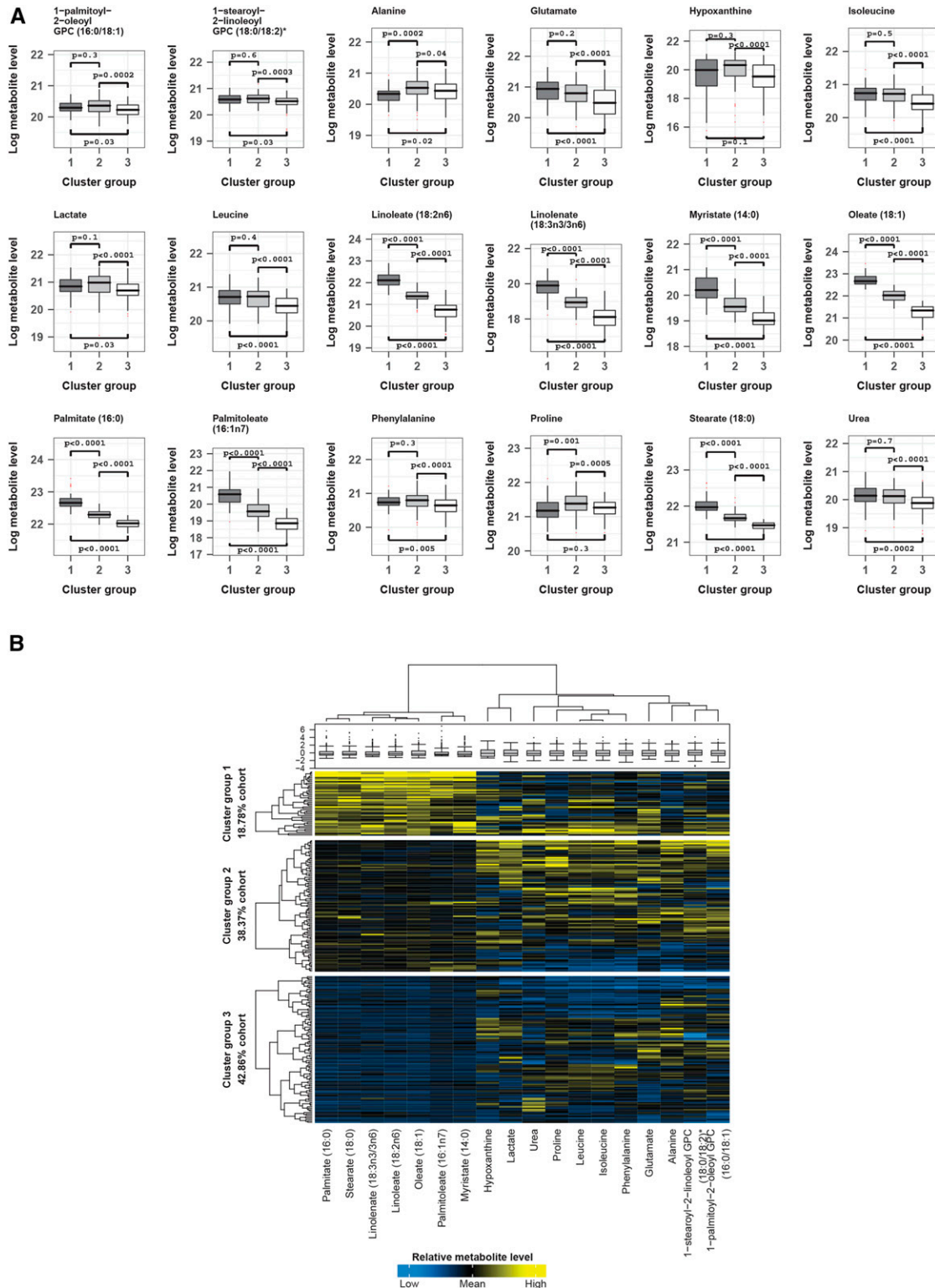
Using the 3-cluster solution, we identified metabolites that were over- or underrepresented in each cluster and those likely to be involved in segregation of each cluster from the other by first centering the raw level medoid vectors for  $k = 3$  together to give a roughly even distribution around a mean = 0 of all medoids. After this simple step, medoids close to the mean will be mostly similar across all cluster groups, whereas those further away will have increasingly stronger influence in determining cluster segregation. To then determine, per cluster, the metabolites responsible for this segregation, we conducted a normalization step that involved scaling each medoid vector (i.e., 3 vectors corresponding to 3 clusters) separately to a constant range of 0 to 1, resulting in those being closer to 0 and 1 (but not in between) having a greater role in driving the segregation of their respective cluster (i.e., they “characterize” the cluster group). The final conversion of these into  $z$  scores provides a standardized scaling system for choosing cutoffs. We used an initial lenient cutoff of  $|z| > 1$  to identify an initial list of metabolites (**Supplemental Figure 3**). The  $\log_e$  concentrations of these metabolites were then compared across clusters, and those with a statistically significant difference (Kruskal-Wallis  $P < 0.001$ )

**TABLE 1**Maternal and infant characteristics per cluster group, ordered as clusters 1–3 based on decreasing concentrations of inflammatory fatty acids and amines<sup>1</sup>

Characteristics	Cluster 1 (n = 46)	Cluster 2 (n = 94)	Cluster 3 (n = 105)	P	$\beta$ -coefficient <sup>2</sup>	OR (95% CI)
<b>Maternal characteristics</b>						
Age at enrollment, y	25.6 $\pm$ 5.37	28.1 $\pm$ 5.26	27.3 $\pm$ 5.51	0.017	0.074	1.077 (1.013, 1.15)
Gestational age at enrollment, d	99.8 $\pm$ 19.3	96.2 $\pm$ 21.5	96.8 $\pm$ 18.9	0.31	0.0083	0.99 (0.98, 1.01)
Education						
Graduate school	3 (6.52)	20 (21.28)	17 (16.19)	0.012	1.91	6.78 (1.87, 32.63)
College graduate	7 (15.22)	18 (19.15)	22 (20.95)		1.15	3.14 (1.076, 9.74)
Junior college or some college	8 (17.39)	26 (27.66)	30 (28.57)		1.35	3.85 (1.37, 11.3)
Technical school	2 (4.35)	4 (4.25)	6 (5.72)		1.01	2.75 (0.59, 20.02)
High school graduate?						
Yes	15 (32.61)	15 (15.96)	21 (20.00)		0.28	1.32 (0.5, 3.42)
No	11 (23.91)	11 (11.70)	9 (8.57)			—
Household income, \$/y						
>150,000	1 (2.17)	4 (4.26)	5 (4.76)	0.2	0.74	2.089 (0.35, 40.25)
100,000–149,999	3 (6.52)	6 (6.38)	14 (13.33)		0.44	1.55 (0.44, 7.25)
75,000–99,999	5 (10.87)	15 (15.96)	9 (8.57)		0.11	1.11 (0.37, 3.78)
50,000–74,999	1 (2.17)	14 (14.89)	7 (6.67)		1.58	4.88 (0.89, 91.27)
30,000–49,999	6 (13.05)	10 (10.64)	17 (16.19)		0.044	1.045 (0.37, 3.25)
Other or unreported	17 (36.96)	19 (20.21)	23 (21.91)		–0.56	0.57 (0.25, 1.3)
<30,000	13 (28.26)	26 (27.66)	30 (28.57)			—
Marital status						
Married	20 (43.48)	47 (50.00)	50 (47.62)	0.34	0.19	1.21 (0.55, 2.61)
Separated or divorced	0 (0)	3 (3.19)	3 (2.86)		15.18	NA
Not married						
Living together	13 (28.26)	23 (24.47)	21 (20.00)		–0.17	0.85 (0.35, 2.03)
Not living together	13 (28.26)	21 (22.34)	31 (29.52)			—
Treatment arm						
Vitamin D	21 (45.65)	47 (50.00)	51 (48.57)	0.66	0.14	1.16 (0.61, 2.21)
Placebo	25 (54.35)	47 (50.00)	54 (51.42)			—
<b>Infant characteristics</b>						
Sex						
Female	26 (56.52)	51 (54.26)	56 (53.33)	0.74	–0.111	0.9 (0.47, 1.7)
Male	20 (43.48)	43 (45.74)	49 (46.67)			—
Ethnicity						
Asian	2 (4.35)	7 (7.45)	5 (4.76)	0.098	0.73	2.07 (0.52, 13.78)
Caucasian	10 (21.74)	32 (34.04)	36 (34.29)		0.85	2.34 (1.1, 5.36)
Native Hawaiian	0 (0.00)	1 (1.06)	2 (1.91)		14.5	NA
Other	4 (8.69)	14 (14.90)	15 (14.28)		0.92	2.5 (0.89, 8.94)
African American	30 (65.22)	40 (42.55)	47 (44.76)			—
Length at birth, cm	50.5 $\pm$ 2.74	51.2 $\pm$ 2.79	51.0 $\pm$ 2.41	0.19	0.082	1.085 (0.96, 1.23)
Weight at birth, kg	3.24 $\pm$ 0.49	3.41 $\pm$ 0.5	3.34 $\pm$ 0.51	0.11	0.00051	1 (1, 1)
BMI, kg/m <sup>2</sup>	12.6 (9.80–15.40)	12.95 (9.69–17.25)	12.78 (8.93–16.41)	0.28	0.13	1.14 (0.9, 1.46)
Mode of delivery						
Cesarean	17 (36.96)	59 (62.77)	34 (32.38)	0.68	–0.18	0.83 (0.43, 1.66)
Vaginal						
With previous cesarean	1 (2.17)	1 (1.06)	0 (0.0)		–1.52	0.22 (0.0084, 5.62)
Forceps assisted	0 (0.0)	0 (0.0)	2 (1.91)		14.042	NA
Vacuum assisted	1 (2.17)	3 (3.19)	4 (3.81)		0.42	1.52 (0.26, 29.1)
Infant still being breastfed						
Month 12	11 (23.91)	25 (26.60)	25 (23.81)	0.18	–1.26	0.28 (0.015, 1.64)
Since last visit (month 24)	34 (73.91)	78 (82.98)	84 (80.00)	0.19	0.52	1.68 (0.79, 3.52)
Since last visit (month 33)	33 (71.74)	75 (79.78)	84 (80.00)	0.19	0.5	1.65 (0.77, 3.38)
Vitamin D at 3 y, ng/mL	16.6 $\pm$ 7.25	21.9 $\pm$ 9.24	22.0 $\pm$ 8.76	0.00008	0.086	1.09 (1.042, 1.14)
<b>Sample characteristics</b>						
Sample storage time, y	5.75 $\pm$ 0.5	5.57 $\pm$ 0.5	5.50 $\pm$ 0.52	0.0097	–0.84	0.43 (0.22, 0.82)

<sup>1</sup> Values are means  $\pm$  SDs or *n* (%) unless otherwise indicated. Characteristics showing evidence of confounding of cluster membership included mother's age at enrollment, mother's educational level, child's vitamin D concentration at age 3 y, and sample storage time, but no statistically significant differences across the clusters with respect to sex, ethnicity, anthropometrics, vaginal compared with cesarean delivery, or breastfeeding patterns. All *P* values were derived by applying  $\chi^2$  ANOVA to each multinomial logistic regression model, with sample-to-cluster assignment as the outcome and each characteristic as the predictor of this. All *P* values,  $\beta$ -coefficients, and ORs are derived from  $\chi^2$  ANOVA on a multinomial logistic regression model with sample-to-cluster assignment as outcome, with levels ordered as clusters 1–3 (i.e., decreasing concentrations of inflammatory fatty acids and amines). Reference categories were mother's education, did not graduate from high school; household income, <\$30,000/y; mother's marital status, not married and not living together; treatment arm, placebo; sex, male; ethnicity, African American; infant still being breastfed (month 12), no; infant breastfed since last visit (month 24), no; infant breastfed since last visit (month 33), no; mode of delivery, vaginal. NA, not applicable.

<sup>2</sup> The  $\beta$ -coefficient is the rate of change of each maternal and infant characteristic per unit change in the outcome of interest.



**FIGURE 1** Log<sub>2</sub> levels of metabolites across cluster groups highlights the main drivers of group segregation. Only metabolites identified as exhibiting variability across the identified cluster groups by medoid  $|z| > 1$  in the 3-cluster solution and that were additionally significant by Kruskal-Wallis  $P < 0.001$  across these groups were chosen. Their natural log metabolite concentrations were plotted to reveal true patterns of similarity and difference across each group. (A) Box plots with pairwise group comparisons for each metabolite (Mann-Whitney  $U$  test). (B) Sample and metabolite cluster dendrograms with heat map and box plot of metabolite concentrations. Sample dendrogram and heat map divided by sample-to-cluster assignment. Total sample  $N = 245$ : group 1, 46; group 2, 94; group 3, 105. GPC, glycerophosphorylcholine.

were defined as over- or underrepresented and used for further analyses to describe the clusters.

### In utero vitamin D exposure and cluster membership

We first examined differences in baseline characteristics across children in the 3 clusters by building multinomial logistic regression models with sample-to-cluster assignment as outcome (i.e., 3 possible outcomes relating to cluster 1, 2, or 3) to identify variables that may be confounding the relation between in utero vitamin D exposure and cluster membership. Second, we investigated the influence of maternal pre- and postsupplement vitamin D concentrations on cluster membership, again using the same multinomial logistic regression model with sample-to-cluster assignment as the outcome but now adjusting for the potential confounders, which were the mother's age at enrollment and education level, child's vitamin D concentration at age 3 y, and sample storage time. Child asthma status at age 3 y was also included as a potential confounder (Table 1). We built a directed acyclic graph (DAG) with DAGitty 2.3 software (23) (<http://dagitty.net/>) to identify the appropriate confounding variables to adjust for (Supplemental Table 1, Supplemental Figure 4). Reported  $P$  values were from  $\chi^2$  ANOVA. Third, we created a combined variable to analyze the influence of vitamin exposure throughout pregnancy by using a cutoff of 30 ng/mL (24) to define low (<30 ng/mL) and normal ( $\geq$ 30 ng/mL) maternal pre- and postsupplement vitamin D concentrations. This resulted in 4 categories: low-low (low presupplement and post-supplement concentrations), low-high, high-low, and high-high. The association between this categoric variable and cluster membership was analyzed with Pearson's  $\chi^2$  test of independence. A

5% significance level was used for all tests. All analyses were conducted in R 3.3.1 software (R Foundation for Statistical Computing, 2016-06-21) (25).

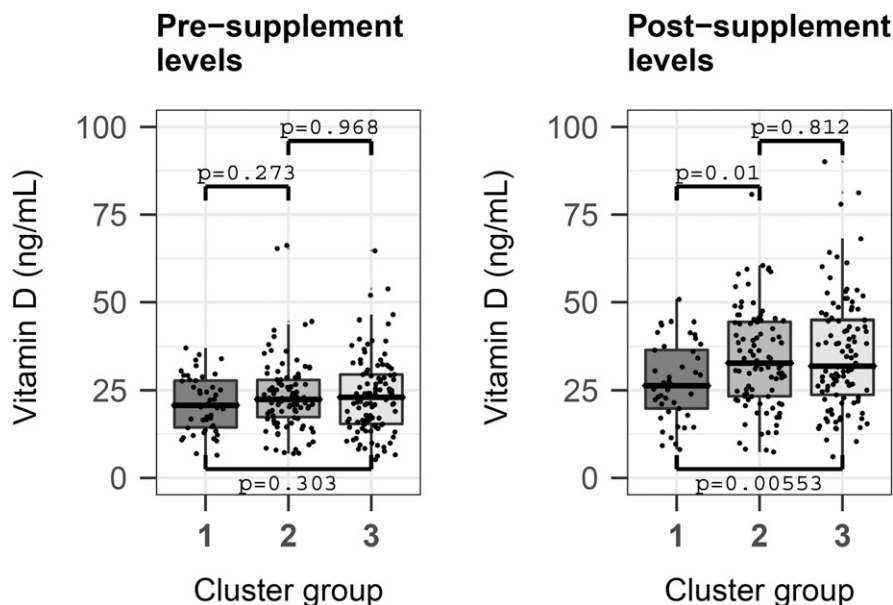
## RESULTS

### Baseline characteristics

The case-control subset selected for metabolomic profiling consisted of 245 children; 84 (34%) had asthma, 112 (46%) were boys, 117 (48%) were African American, 78 (32%) were Caucasian, and the remaining children were of Asian, Hawaiian, or other ethnic origin. A summary of the baseline characteristics of the selected population by cluster membership is provided in Table 1, showing a statistically significant influence on cluster membership by mother's age at enrollment, mother's education level, child's vitamin D concentration at age 3 y, and sample storage time, but no statistically significant differences across the clusters with respect to sex, ethnicity, anthropometrics, vaginal compared with cesarean delivery, or breastfeeding patterns. A DAG illustrating our causal pathway and statistically significant confounders is provided in Supplemental Figure 4.

### Clusters of children based on metabolomic profiles

The children were divided into 3 distinct data-driven clusters by their metabolomic profiles, with 46 (19%) children in cluster 1, 94 (38%) in cluster 2, and 105 (43%) in cluster 3. The concentrations of most of the metabolites did not differ between each cluster, but the differences between the 3 clusters for 18 metabolites were statistically significant ( $P < 0.001$ ) (Figure 1A).



**FIGURE 2** Pre- and postsupplement maternal vitamin D concentrations significantly associate with cluster group segregation. Postsupplement concentrations were significantly lower when cluster group 1 was compared with groups 2 and 3 separately (Mann-Whitney  $U$  test). In addition, through a multinomial logistic regression model with sample-to-cluster assignment as outcome and adjusting for significant confounders, sample storage time, maternal age, and education, and the child's asthma status and vitamin D level at age 3 y, the postsupplement concentration of vitamin D was a significant predictor of cluster grouping ( $P = 0.0014$ ; OR: 1.032; 95% CI: 1.0021, 1.065). The presupplementation concentration of vitamin D was not a statistically significant predictor of cluster grouping. Presupplementation concentrations (means  $\pm$  SDs): group 1, 21.10  $\pm$  8.01 ng/mL; group 2, 23.60  $\pm$  10.30 ng/mL; and group 3, 23.60  $\pm$  11.00 ng/mL. Postsupplementation concentrations (means  $\pm$  SDs): group 1, 27.5  $\pm$  11.0 ng/mL; group 2, 34.0  $\pm$  14.1 ng/mL; and group 3 (35.2  $\pm$  15.9 ng/mL). Total sample  $N = 245$ : group 1, 46; group 2, 94; and group 3, 105.

Fatty acids, such as linoleate (18:2n-6), oleate (18:1), palmitate (16:0), and myristate (14:0), among others, were the main drivers of cluster segregation, with the highest amounts in cluster 1, intermediate amounts in cluster 2, and lowest amounts in cluster 3. Amounts of amines and their derivatives, which included glutamate, hypoxanthine, isoleucine, lactate, leucine, phenylalanine, and urea, were comparable across cluster groups 1 and 2, and were statistically significantly higher in either or both of these groups compared with cluster group 3. Figure 1B shows a clustering dendrogram and heat map of metabolite concentrations across the clusters, illustrating how particular fatty acids and amines were responsible for cluster segregation.

### In utero vitamin D exposure and cluster segregation

Maternal postsupplement vitamin D concentrations were statistically significantly associated with the child's cluster membership at age 3 y, with lower concentrations in cluster 1 (mean  $\pm$  SD: 27.5  $\pm$  11.0 ng/mL) compared with cluster 2 (34.0  $\pm$  14.1 ng/mL) and cluster 3 (35.2  $\pm$  15.9 ng/mL) (overall  $P = 0.0018$ ; OR: 1.04; 95% CI: 1.014, 1.07). This relation was robust to adjustment for sample storage time, maternal age and education, and the child's asthma status and vitamin D concentration at age 3 y ( $P = 0.0014$ ; OR: 1.032; 95% CI: 1.0021, 1.065). There was no statistically significant difference in maternal presupplement vitamin D concentrations in cluster 1 compared with 2 and 3: 21.10  $\pm$  8.01 compared with 23.60  $\pm$  10.30 compared with 23.60  $\pm$  11.00 ng/mL ( $P = 0.12$ ; OR: 1.03; 95% CI: 0.99, 1.06; adjusted model  $P = 0.116$ ; OR: 0.99; 95% CI: 0.96, 1.04) (Figure 2).

The combined variable of maternal pre- and postsupplement vitamin D concentrations of <30 ("low") compared with  $\geq 30$  ng/mL ("high") showed a distribution of 40.2% with low-low, 40.2% with low-high, 5.7% with high-low, and 13.9% with high-high. This combined variable illustrating in utero vitamin D exposure throughout pregnancy was a statistically significant determinant of the child's cluster membership ( $P = 0.03$ ), with an overrepresentation of low-low and underrepresentation of high-high in cluster 1, suggesting that the presupplement concentrations combined with postsupplement concentrations play a role for cluster segregation (Figure 3).

### DISCUSSION

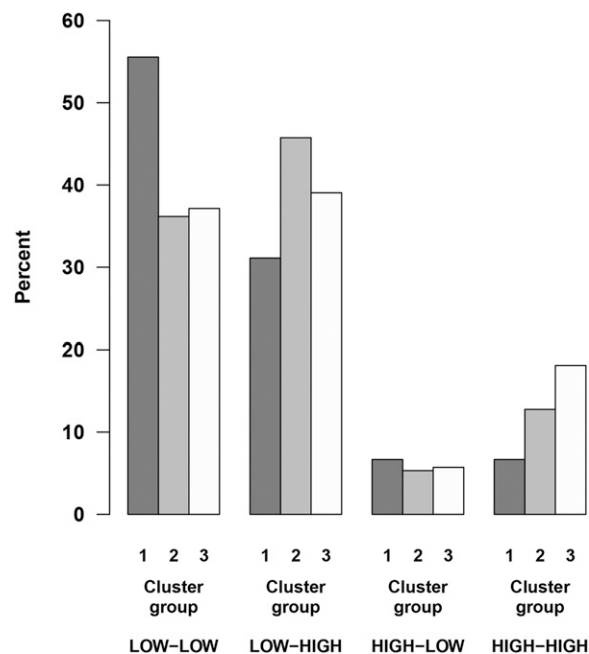
We determined that a subset of 3-y-old children from the VDAART cohort could be segregated into 3 distinct clusters by their plasma metabolomics profiles. These clusters were mainly driven by differences in concentrations of inflammatory fatty acids and amines and amine derivatives. Cluster membership was statistically significantly associated with maternal post-supplementation vitamin D concentrations and the child's vitamin D concentration at age 3 y, suggesting a strong influence of vitamin D on the metabolome and alluding to a prenatal programming role of vitamin D on the child's metabolism.

High amounts of fatty acids, in particular linoleate, linolenate (18:3n-3/3n-6), myristate, oleate, palmitate, palmitoleate (16:1n-7), and stearate (18:0), characterized children in cluster 1. Compared with clusters 2 and 3, the children in cluster 1 were exposed to the lowest in utero vitamin D concentrations throughout pregnancy, with significantly lower postsupplement concentrations at weeks 32-38 of gestation, and the highest proportion of children

exposed to concentrations <30 ng/mL both pre- and post-supplement (low-low). The metabolomic profile of the children in cluster 3 represented the reverse of cluster 1, exhibiting the lowest concentrations of the aforementioned fatty acids and was associated with normal concentrations of vitamin D. It is biologically interesting that cluster membership was influenced by in utero vitamin D exposure and was mainly determined by concentrations of fatty acids, because unfavorable fatty acid composition can promote inflammation and increase risk of complex diseases (26, 27).

Some of the fatty acids driving cluster segregation were the SFAs palmitate and myristate, which had highest concentrations in cluster 1, intermediate concentrations in cluster 2, and lowest concentrations in cluster 3 (i.e., highest concentrations in the cluster with lowest in utero vitamin D exposure). Both of these SFAs have been shown to have proinflammatory disease-promoting capacities. Increased intake of palmitate, found in butter, cheese, and meat, for example, has been associated with cardiovascular disease and inflammation (28, 29), whereas increased intake of myristate, which is found in butter and many other animal fats, has been associated with elevated inflammatory markers such as C-reactive protein and several other proinflammatory cytokines (30). In general, increased intake of saturated fats can lead to dyslipidemia, which has recently been linked to childhood asthma and diminished lung function (31).

Another important fatty acid responsible for cluster segregation was the omega-6 PUFA linoleate, showing highest concentrations



**FIGURE 3** The combination of maternal vitamin D concentrations pre- and postsupplement is a determinant of cluster membership. Maternal pre- and postsupplement vitamin D concentrations were encoded as "low" (<30 ng/mL) or "high" ( $\geq 30$  ng/mL), forming the basis for a combined variable describing the intrauterine exposure at both times (e.g., low-low means that both pre- and postsupplement concentrations of vitamin D were <30 ng/mL). This combined variable was a strong determinant of the child's cluster membership ( $\chi^2 P = 0.03$ ). An overrepresentation of low-low and underrepresentation of high-high is observed in cluster 1, suggesting that the presupplement concentrations combined with postsupplement concentrations play a role for cluster segregation. Total sample  $N = 245$ : group 1, 46; group 2, 94; and group 3, 105.

in cluster 1, which was associated with low in utero vitamin D exposure. Linoleate is a precursor of arachidonate, prostaglandins, and leukotrienes, which are known components of the inflammatory cascade and involved in asthmatic airway inflammation (32). Linoleate, along with oleate, which is a monounsaturated  $\omega$ -9 fatty acid and is the main component of olive oil, has also been shown to have anti-inflammatory properties by reducing macrophage numbers in hypercholesterolemia (33) and reducing concentrations of plasminogen activator inhibitor-1, which is a marker of inflammation (34). Thus, linoleate and other fatty acids be both pro- and anti-inflammatory, which is believed to be determined by the balance of several PUFAs from the diet (35) that compete for the same rate-limiting enzymes either promoting or resolving inflammation (36). Furthermore, other dietary micronutrients, such as vitamin D, may be important in determining the health effects of fatty acids, which is supported by a recent prenatal PUFA supplement trial (37) showing a reduced risk of offspring asthma that was dependent on concurrent vitamin D supplement status. Our study shows that low in utero vitamin D exposure is associated with membership of a metabolic cluster characterized by high concentrations of the  $\omega$ -6 PUFA linoleate, whereas sufficient exposure determines membership of metabolic clusters with lower concentrations of  $\omega$ -6 PUFAs.

Concentrations of a range of amines were also responsible for cluster segregation. In particular, children in clusters 1 and 2, who were exposed to low and normal concentrations of vitamin D in utero, respectively, had a common metabolomic profile characterized by high concentrations of amines, such as alanine, glutamate, hypoxanthine, lactate, phenylalanine, and proline. Cluster group 2 differed from group 1, however, through concentrations of inflammatory fatty acids already mentioned (higher in group 1). Some of the identified amines, and amines broadly, are known inflammatory biomarkers, such as hypoxanthine and glutamate in arthritis (38, 39), and arginine has more recently been implicated in asthma as a result of its role in bioenergetics (40). Overall, these findings suggest that the children in cluster 2 who go on to develop inflammation and complex diseases later in life may do so through perturbations of amino acids pathways rather than fatty acid metabolism.

A major strength of this study is the state-of-the-art UPLC-MS/MS metabolomic profiling platform with metabolite identification that used known standards. This resulted in a data set with a high number of annotated metabolites, which enabled a meaningful biological interpretation of our results. Still, as technology improves, the increasing number of metabolites that may be identified could add further knowledge about the potential prenatal programming effect of vitamin D on the child's metabolism. The implementation of various clustering metrics, such as the gap statistic and silhouette coefficient before the clustering of the cohort, eliminated the need for a priori decisions on the number of clusters. Furthermore, the clustering technique was purely data driven and independent of any assumptions on exposure or endpoints associations. This analytic approach avoids the classical issue of overfitted models trained on the outcome or exposure of interest.

Another major advantage is the mother-child cohort design with multiple measures of vitamin D status during pregnancy, which is a strength compared with a single assessment that can be subject to multiple confounders. Furthermore, the embedded vitamin D intervention trial (18) presumably resulted in a perturbation of vitamin D-related metabolic pathways, providing

us the power to discover differences, which may be subtle and less easy to discover in the normal range of in utero vitamin D exposure. However, the high-risk setting of the VDAART cohort, where all mothers or their partners have  $\geq 1$  atopic disorder, is a potential limitation of the study that may hamper generalization of the result to unselected populations. Also, the plasma metabolomics profiling was done on a nested case-control group with an overrepresentation of asthmatic children.

Our results suggest a link between both in utero and childhood vitamin D exposure and the metabolism of the child at age 3 y. Nevertheless, to implicate true causality from our study's results is not possible at this point, even though our analysis included exhaustive measures to exclude the influence of other characteristics on sample segregation into cluster groups and the differing concentrations of metabolites across these.

In conclusion, this study shows that 3-y-old children can be segregated into 3 distinct biologically meaningful clusters by their plasma metabolomic profiles. Cluster membership was influenced by in utero and childhood vitamin D exposure and was predominantly determined by fatty acids and amines with differing patterns and concentrations across the clusters. These findings suggest that vitamin D concentrations influence the metabolome via key metabolites related to inflammatory pathways and allude to a prenatal programming role of vitamin D on the child's metabolism at age 3 y, which may be of importance for susceptibility to inflammation and complex diseases later in life.

The authors' responsibilities were as follows—KB: analyzed all data, performed all statistical analyses, and jointly wrote the manuscript with BLC; KB; conceived the initial research plan, with significant later contributions from BLC for this study; KB, BLC, RSK, and JAL-S: interpreted the results and contributed to the overall intellectual design and evolution of the study; MM and HM: assisted in the elucidation of the methods; AAL and STW: designed the VDAART clinical trial; and all authors: read and approved the final manuscript. JAL-S is a consultant to Metabolon Inc. (North Carolina). The remaining authors reported no conflicts of interest related to the study.

## REFERENCES

- Müller MJ, Volmer DA. Mass spectrometric profiling of vitamin D metabolites beyond 25-hydroxyvitamin D. *Clin Chem* 2015;61:1033–48.
- Holick MF. Vitamin D deficiency. *N Engl J Med* 2007;357:266–81.
- Ginde AA, Mansbach JM, Camargo CA. Vitamin D, respiratory infections, and asthma. *Curr Allergy Asthma Rep* 2009;9:81–7.
- Pilz S, Verheyen N, Grubler MR, Tomaschitz A, Marz W. Vitamin D and cardiovascular disease prevention. *Nat Rev Cardiol* 2016;13:404–17.
- Deeb KK, Trump DL, Johnson CS. Vitamin D signalling pathways in cancer: potential for anticancer therapeutics. *Nat Rev Cancer* 2007;7:684–700.
- Spedding S. Vitamin D and depression: a systematic review and meta-analysis comparing studies with and without biological flaws. *Nutrients* 2014;6:1501–18.
- Jeffery LE, Raza K, Hewison M. Vitamin D in rheumatoid arthritis—towards clinical application. *Nat Rev Rheumatol* 2016;12:201–10.
- Abrahamsen B, Harvey NC. The role of vitamin D supplementation in patients with rheumatic diseases. *Nat Rev Rheumatol* 2013;9:411–22.
- Agmon-Levin N, Theodor E, Segal RM, Shoenfeld Y. Vitamin D in systemic and organ-specific autoimmune diseases. *Clin Rev Allergy Immunol* 2013;45:256–66.
- Mangin M, Sinha R, Fincher K. Inflammation and vitamin D: the infection connection. *Inflamm Res* 2014;63:803–19.
- Nair R, Maseeh A. Vitamin D: the “sunshine” vitamin. *J Pharmacol Pharmacother* 2012;3:118–26.
- Braithwaite MC, Kumar P, Tyagi C, Tomar LK, Choonara YE, Pillay V. Vitamin D therapy and related metabolomics: is the calciferol dose and form the only requirements for successful clinical therapeutics? *Med Hypotheses* 2013;81:656–63.

13. Eder W, Ege MJ, von Mutius E. The asthma epidemic. *N Engl J Med* 2006;355:2226–35.
14. Mirzakhani H, Al-Garawi A, Weiss ST, Litonjua AA. Vitamin D and the development of allergic disease: how important is it? *Clin Exp Allergy* 2015;45:114–25.
15. Cantorna MT, Snyder L, Lin YD, Yang L. Vitamin D and 1,25(OH)<sub>2</sub>D regulation of T cells. *Nutrients* 2015;7:3011–21.
16. Aranow C. Vitamin D and the immune system. *J Investig Med* 2011;59:881–6.
17. Kelly RS, Dahlin A, McGeachie MJ, Qiu W, Sordillo J, Wan ES, Wu AC, Lasky-Su J. Asthma metabolomics and the potential for integrative omics in research and the clinic. *Chest* 2017;151:262–77.
18. Litonjua AA, Lange NE, Carey VJ, Brown S, Laranjo N, Harshfield BJ, O'Connor GT, Sandel M, Strunk RC, Bacharier LB, et al. The Vitamin D Antenatal Asthma Reduction Trial (VDAART): rationale, design, and methods of a randomized, controlled trial of vitamin D supplementation in pregnancy for the primary prevention of asthma and allergies in children. *Contemp Clin Trials* 2014;38:37–50.
19. Litonjua AA, Carey VJ, Laranjo N, Harshfield BJ, McElrath TF, O'Connor GT, Sandel M, Iverson RE Jr., Lee-Paritz A, Strunk RC, et al. Effect of prenatal supplementation with vitamin D on asthma or recurrent wheezing in offspring by age 3 years: the VDAART randomized clinical trial. *JAMA* 2016;315:362–70.
20. Ersfeld DL, Rao DS, Body JJ, Sackrison JL Jr., Miller AB, Parikh N, Eskridge TL, Polinske A, Olson GT, MacFarlane GD. Analytical and clinical validation of the 25 OH vitamin D assay for the LIAISON automated analyzer. *Clin Biochem* 2004;37:867–74.
21. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc Series B Stat Methodol* 2001;63:411–23.
22. Rousseeuw P. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987;20:53–65.
23. Textor J, Hardt J, Knüppel S. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology* 2011;22:745.
24. Mirzakhani H, Litonjua AA, McElrath TF, O'Connor G, Lee-Parritz A, Iverson R, Macones G, Strunk RC, Bacharier LB, Zeiger R, et al. Early pregnancy vitamin D status and risk of preeclampsia. *J Clin Invest* 2016;126:4702–15.
25. R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2016 [cited 2017 Apr 6]. Available from: <https://www.R-project.org/>.
26. Lands WE, Libelt B, Morris A, Kramer NC, Prewitt TE, Bowen P, Schmeisser D, Davidson MH, Burns JH. Maintenance of lower proportions of (n - 6) eicosanoid precursors in phospholipids of human plasma in response to added dietary (n - 3) fatty acids. *Biochim Biophys Acta* 1992;1180:147–62.
27. Lee TH, Mencia-Huerta JM, Shih C, Corey EJ, Lewis RA, Austen KF. Effects of exogenous arachidonic, eicosapentaenoic, and docosahexaenoic acids on the generation of 5-lipoxygenase pathway products by ionophore-activated human neutrophils. *J Clin Invest* 1984;74:1922–33.
28. Sokolova M, Vinge LE, Alfsnes K, Olsen MB, Eide L, Kaasbøll OJ, Attramadal H, Torp MK, Fosshaug LE, Rashidi A, et al. Palmitate promotes inflammatory responses and cellular senescence in cardiac fibroblasts. *Biochim Biophys Acta* 2017;1862:234–45.
29. Pillon NJ, Azizi PM, Li YE, Liu J, Wang C, Chan KL, Hopperton KE, Bazinet RP, Heit B, Bilan PJ, et al. Palmitate-induced inflammatory pathways in human adipose microvascular endothelial cells promote monocyte adhesion and impair insulin transcytosis. *Am J Physiol Endocrinol Metab* 2015;309:E35–44.
30. Perreault M, Roke K, Badawi A, Nielsen DE, Abdelmagid SA, El-Sohemy A, Ma DW, Mutch DM. Plasma levels of 14:0, 16:0, 16:1n-7, and 20:3n-6 are positively associated, but 18:0 and 18:2n-6 are inversely associated with markers of inflammation in young healthy adults. *Lipids* 2014;49:255–63.
31. Vinding RK, Stokholm J, Chawes BL, Bisgaard H. Blood lipid levels associate with childhood asthma, airway obstruction, bronchial hyperresponsiveness, and aeroallergen sensitization. *J Allergy Clin Immunol* 2016;137:68–74.e4.
32. Weiss JW, Drazen JM, McFadden E Jr., Weller P, Corey EJ, Lewis RA, Austen KF. Airway constriction in normal humans produced by inhalation of leukotriene D: potency, time course, and effect of aspirin therapy. *JAMA* 1983;249:2814–7.
33. Reaven P, Parthasarathy S, Grasse BJ, Miller E, Steinberg D, Witztum JL. Effects of oleate-rich and linoleate-rich diets on the susceptibility of low density lipoprotein to oxidative modification in mildly hypercholesterolemic subjects. *J Clin Invest* 1993;91:668–76.
34. Silver HJ, Kang H, Keil CD, Muldowney JA III, Kocalis H, Fazio S, Vaughan DE, Niswender KD. Consuming a balanced high fat diet for 16 weeks improves body composition, inflammation and vascular function parameters in obese premenopausal women. *Metabolism* 2014;63:562–73.
35. McDade TW. Early environments and the ecology of inflammation. *Proc Natl Acad Sci U S A* 2012;109 Suppl 2:17281–8.
36. Serhan CN. Pro-resolving lipid mediators are leads for resolution physiology. *Nature* 2014;510:92–101.
37. Bisgaard H, Stokholm J, Chawes BL, Vissing NH, Bjarnadóttir E, Schoos AM, Wolsk HM, Pedersen TM, Vinding RK, Thorsteinsdóttir S, et al. Fish oil-derived fatty acids in pregnancy and wheeze and asthma in offspring. *N Engl J Med* 2016;375:2530–9.
38. Gudbjörnsson B, Zak A, Niklasson F, Hällgren R. Hypoxanthine, xanthine, and urate in synovial fluid from patients with inflammatory arthritides. *Ann Rheum Dis* 1991;50:669–72.
39. Hajati AK, Alstergren P, Näsström K, Bratt J, Kopp S. Endogenous glutamate in association with inflammatory and hormonal factors modulates bone tissue resorption of the temporomandibular joint in patients with early rheumatoid arthritis. *J Oral Maxillofac Surg* 2009;67:1895–903.
40. Xu W, Ghosh S, Comhair SA, Asosingh K, Janocha AJ, Mavrakis DA, Bennett CD, Gruca LL, Graham BB, Queisser KA, et al. Increased mitochondrial arginine metabolism supports bioenergetics in asthma. *J Clin Invest* 2016;126:2465–81.