



# Dried blood spot metabolomics reveals a metabolic fingerprint with diagnostic potential for Diamond Blackfan Anaemia

Birgit van Dooijeweert,<sup>1,2,\*</sup>   
 Melissa H. Broeks,<sup>3,\*</sup> Eduard J. van  
 Beers,<sup>4</sup>  Nanda M. Verhoeven-Duif,<sup>3</sup>  
 Wouter W. van Solinge,<sup>1</sup> Edward E.  
 S. Nieuwenhuis,<sup>2</sup> Judith J. Jans,<sup>3</sup>  
 Richard van Wijk<sup>1,\*</sup> and  
 Marije Bartels<sup>2,4,\*</sup>

<sup>1</sup>Central Diagnostic Laboratory-Research, University Medical Center Utrecht, Utrecht University, <sup>2</sup>Department of Paediatric Haematology, University Medical Center Utrecht, Utrecht University,

<sup>3</sup>Section Metabolic Diagnostics, Department of Genetics, University Medical Center Utrecht, Utrecht University, and <sup>4</sup>Van Creveldkliniek, University Medical Center Utrecht, Utrecht, the Netherlands

Received 24 February 2021; accepted for publication 18 March 2021

Correspondence: Birgit van Dooijeweert, Central Diagnostic Laboratory-Research, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX, Utrecht, the Netherlands.

Email: b.vandooijeweert@umcutrecht.nl

\*Authors contributed equally.

## Introduction

Diamond Blackfan Anaemia (DBA, OMIM# 105650) is a rare inherited bone marrow failure syndrome (IBMFS) characterised by erythroid hypoplasia, congenital malformations (~50%), growth defects and an increased risk of developing malignancies.<sup>1–3</sup> From a clinical and genetic perspective the disorder is highly heterogeneous, and clear genotype-phenotype correlations are absent. Since the majority of molecular defects have been found in ribosomal protein (RP) genes, resulting in impaired ribosome biogenesis, DBA is regarded a ‘ribosomopathy’.<sup>4,5</sup>

While its genetic basis has been studied intensively, the pathophysiology of DBA is still not fully understood. One of the major unresolved issues is how RP gene mutations result in the specific erythroid defect that characterises DBA.<sup>6</sup>

## Summary

The diagnostic evaluation of Diamond Blackfan Anaemia (DBA), an inherited bone marrow failure syndrome characterised by erythroid hypoplasia, is challenging because of a broad phenotypic variability and the lack of functional screening tests. In this study, we explored the potential of untargeted metabolomics to diagnose DBA. In dried blood spot samples from 18 DBA patients and 40 healthy controls, a total of 1752 unique metabolite features were identified. This metabolic fingerprint was incorporated into a machine-learning algorithm, and a binary classification model was constructed using a training set. The model showed high performance characteristics (average accuracy 91.9%), and correct prediction of class was observed for all controls ( $n = 12$ ) and all but one patient ( $n = 4/5$ ) from the validation or ‘test’ set (accuracy 94%). Importantly, in patients with congenital dyserythropoietic anaemia (CDA) – an erythroid disorder with overlapping features – we observed a distinct metabolic profile, indicating the disease specificity of the DBA fingerprint and underlining its diagnostic potential. Furthermore, when exploring phenotypic heterogeneity, DBA treatment subgroups yielded discrete differences in metabolic profiles, which could hold future potential in understanding therapy responses. Our data demonstrate that untargeted metabolomics in dried blood spots is a promising new diagnostic tool for DBA.

**Keywords:** untargeted metabolomics, Diamond Blackfan Anaemia, disease fingerprint, dried blood spots, machine-learning algorithm.

Establishing the diagnosis of DBA can be particularly difficult since, in contrast to other IBMFS, no validated functional screening tests exist. Furthermore, the clinical presentation is highly heterogeneous, even within families who share a molecular defect.<sup>7,8</sup>

Currently, the diagnosis relies on consensus criteria and exclusion of other IBMFS and causes of anaemia. The consensus criteria include age (i.e. presentation in the first year of life), macrocytic anaemia, reticulocytopenia and normal marrow cellularity with a paucity of erythroid precursors. Alternative criteria that confirm or support the diagnosis include the identification of a known molecular defect, elevated erythrocyte adenosine deaminase activity and elevated levels of foetal haemoglobin.<sup>9</sup> In order to improve the diagnostic evaluation of DBA as well as our understanding of phenotypic heterogeneity and clinical severity, novel

functional approaches are needed. One such tool might be metabolomics – the large scale, unbiased study of metabolites which directly reflects the biochemical activity and state of a sample – and thereby represents the cellular phenotype.<sup>10</sup>

$$Z - \text{score} = \frac{(\text{Mass peak intensity of Pt or HC sample} - \text{Mean mass peak intensities of metabolic control samples})}{\text{Standard deviation mass peak intensities of metabolic control samples}^\ddagger}$$

Here we identify and report for the first time on a metabolic fingerprint for DBA using untargeted metabolomics in dried blood spots, and demonstrate the potential of this approach in the diagnostic evaluation of DBA.

## Methods (and/or materials)

### *Patients and samples*

Eighteen patients diagnosed with DBA were included. Diagnosis in all patients was based on a combination of the widely used consensus and supporting criteria and/or confirmed defects in DBA-associated genes.<sup>5,9</sup> In addition, six patients diagnosed with the related disorder of congenital dyserythropoietic anaemia (CDA) were included for subgroup analysis. Healthy volunteers (from an institutional blood donor service) served as healthy controls (HC). All patients or their legal guardians approved the use of their remnant samples for method development and validation, in agreement with institutional and national regulations. All procedures followed were in accordance with the ethical standards of the University Medical Center Utrecht and with the Helsinki Declaration of 1976, as revised in 2000. Sampling was performed at least three weeks after the last transfusion in transfused patients. For dried blood spots (DBS), 50 µl aliquots were spotted onto Guthrie card filter paper (Whatman 903 Protein Saver TM cards). All papers were left to dry for at least four hours at room temperature, and were subsequently stored at –80°C in a foil bag with a desiccant package, pending further analysis.

### *Metabolic phenotyping*

Sample preparation, direct infusion high resolution mass spectrometry (DI-HRMS) and data processing were performed as previously reported.<sup>11–13</sup> Mass peak intensities for metabolite annotations were averaged over technical triplicates. In addition, as DI-HRMS is unable to separate isomers, mass peak intensities consisted of the summed intensities of these isomers. Metabolite annotation was performed using a peak-calling bioinformatics pipeline developed in R-programming software, based on the human metabolome database (version 3.6) (<https://github.com/UMCUGenetics/DIMS>).

DBS samples were distributed over several DI-HRMS runs. In each run, an extra set of control samples was included. To

compare the metabolic profiles between DBA, CDA and HC, mass peak intensities for each identified feature were converted to Z-scores. These scores, based on the extra control samples, were calculated by the following formula:

### *Data analysis*

Z-scores calculated from multiple DI-HRMS runs were combined in a final dataset. Data analyses were conducted in MetaboAnalyst without further data filtering or normalisation.<sup>14</sup> Outlying metabolite features were identified using the PCA loadings plot. In total, 15 outlying features were removed, culminating in a final dataset of 1765 unique features corresponding to 3541 metabolite annotations.

Classification of data was performed in R-software (Version 3.6.1) using the ‘*caret*’ package. This package contains a set of data-processing functions that facilitate the generation of classification and regression models. In this study, a support vector machine (SVM) with a linear kernel – the simplest kernel function without further data transformation – was used for classifying DBA and HC samples<sup>¥</sup>. SVM is a supervised machine-learning model that classifies samples based on the mapping of all data into a high dimensional space, allowing for the separation of two groups of samples into distinct regions by the identification of ‘support vectors’. Classification is then achieved by identifying a separating hyperplane – or decision boundary – between support vectors, and projecting new/unclassified samples into this space. For this analysis, additional features with missing values were removed, resulting in 1737 unique features for modelling. Data and R-code are available upon request.

## Results

### *Explorative untargeted metabolomics identifies metabolic fingerprint*

In total, 1752 unique metabolite features were analysed from DBS samples of 18 patients and 40 controls. Baseline characteristics of patients are summarised in Table I. The DBA patient cohort was characterised by a female predominance (72%) and a lower median age than the controls (9-19 vs. 33-8 years). To determine the variation and distinction

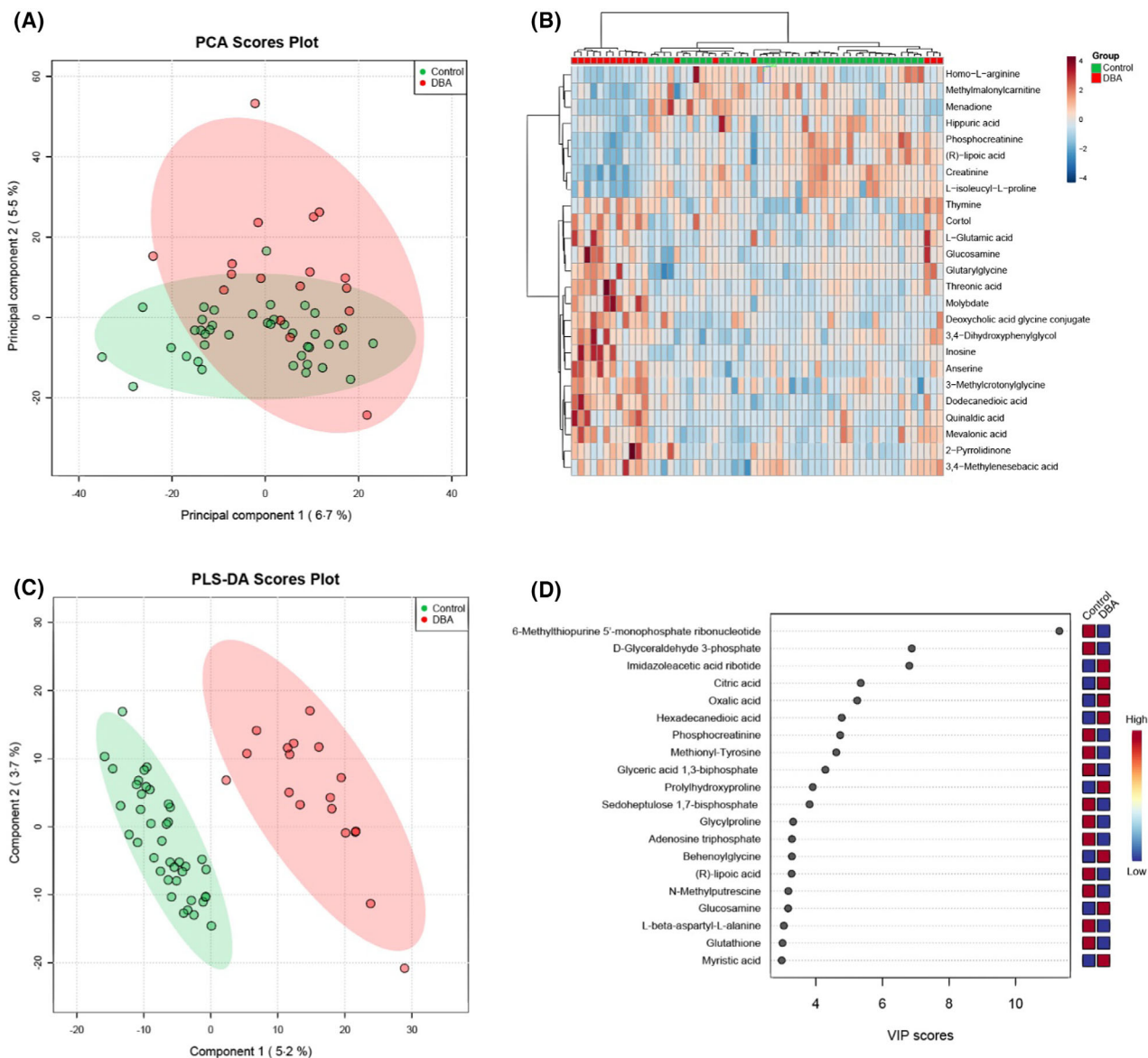
<sup>‡</sup>Metabolic controls exist of a batch of banked DBS samples from individuals in whom an inborn error of metabolism (IEM) was excluded after an extensive diagnostic workup.

<sup>¥</sup>The function of a kernel is to take data as input and transform it into the required form, for example a linear or polynomial kernel.

**Table 1.** (A) Clinical characteristics of DBA patients: age, gender, haemoglobin (Hb), red blood cell count (RBC), reticulocyte count (Retics), mean corpuscular volume (MCV), white blood cell count (WBC), platelets (Plts), treatment and molecular defect. ND, not determined. The patient who was classified incorrectly as a control in the machine-learning model (Fig 2A) is shown in italics. (B) Comparison of age, Hb, Retics and time between blood withdrawal and spotting (time to DBS) between healthy controls (HC) and DBA patients.

| Age (years)   | Gender                     | Hb (mmol/l) | RBC ( $\times 10^{12}/l$ ) | Retics ( $\times 10^9/l$ ) | MCV (fl) | WBC ( $\times 10^9/l$ ) | Plts ( $\times 10^9/l$ ) | Treatment                   | Molecular defect                   |
|---------------|----------------------------|-------------|----------------------------|----------------------------|----------|-------------------------|--------------------------|-----------------------------|------------------------------------|
| 0.7           | Female                     | 5.4         | 2.72                       | 50.8                       | 93.1     | 9.5                     | 508                      | Transfusions                | RPL11; c.396+1G>T; p.(?)           |
| 16            | Female                     | 4.1         | 1.98                       | 61.7                       | 100      | 3.8                     | 260                      | Prednisone (0.17 mg/kg/day) | RPL5; c.493G>T; p.(Gly165*)        |
| 4             | Female                     | 6.2         | 2.87                       | 15.7                       | 105      | 14.0                    | 633                      | Transfusions                | RPS26; c.-362_*del                 |
| 10            | Male                       | 5.0         | 2.53                       | 69.4                       | 100      | 7.1                     | 323                      | Prednisone (0.18 mg/kg/day) | GATA1; c.220+2T>C; p.(?)           |
| 14            | Female                     | 8.0         | 3.49                       | 86.3                       | 109      | 7.3                     | 272                      | Prednisone (0.06 mg/kg/day) | RPS26; c.95_98dup; p.(Asp33GluI6*) |
| 5             | Male                       | 4.1         | 1.90                       | 29.5                       | 92       | 5.0                     | 192                      | Prednisone (0.5 mg/kg/day)  | None identified                    |
| 18            | Female                     | 5.4         | 2.91                       | 17.3                       | 86       | 4.4                     | 294                      | Transfusions                | RPS26; c.2T>C; p.(Met1Thr)         |
| 17            | Female                     | 6.7         | 3.16                       | 33.0                       | 102      | 3.7                     | 183                      | Prednisone (0.04 mg/kg/day) | None identified                    |
| 11            | Female                     | 7.6         | 3.83                       | 48.3                       | 97       | 5.9                     | 337                      | Prednisone (0.07 mg/kg/day) | RPS24; c.*20-2A>G                  |
| 5             | Female                     | 7.3         | 3.61                       | 56.8                       | 93       | 6.7                     | 379                      | None                        | RPS26; c.344T>C; p.(Met115Thr)     |
| 2             | Female                     | 4.3         | 2.31                       | 5.9                        | 86       | 8.6                     | 262                      | Transfusions                | RPL5; c.353C>T; p.(Arg179X)        |
| 6             | Male                       | 7.0         | 3.84                       | 52.3                       | 86       | 19.0                    | 465                      | None                        | RPL9; c.-2+1G>C; p.(?)             |
| 4             | Female                     | ND          | ND                         | ND                         | ND       | ND                      | ND                       | Prednisone (0.5 mg/kg/day)  | RPS26; c.3+1G>C; p.(?)             |
| 5             | Male                       | ND          | ND                         | ND                         | ND       | ND                      | ND                       | Transfusions                | None identified                    |
| 1             | Female                     | 6.5         | 3.77                       | 70.8                       | 86       | 11.9                    | 438                      | None                        | 1p22.1 deletion (incl. RPL5)       |
| 0.8           | Female                     | 4.5         | 2.43                       | 50.1                       | 89       | 8.7                     | 240                      | Transfusions                | None identified                    |
| 21            | Male                       | 9.6         | 4.63                       | 58.8                       | 102      | 3.9                     | 167                      | None                        | 15q25.2 deletion (incl. RPS17)     |
| 25            | Female                     | 6.9         | 3.32                       | 81.5                       | 101      | 4.6                     | 248                      | None                        | RPS19; c.167G>C; p.(Arg56Pro)      |
| Normal range* |                            | 7.4–10.7    | 3.6–5.5                    | 25–120                     | 70–97    | 4.0–13.5                | 150–450                  |                             |                                    |
|               |                            | DBA         |                            |                            |          |                         |                          | HC                          |                                    |
| (B)           | Age (years)                |             |                            | 9.19 ± 7.59                |          |                         |                          |                             | 33.8 ± 8.64                        |
|               | Hb (mmol/l)                |             |                            | 6.16 ± 1.58                |          |                         |                          |                             | 9.07 ± 0.80                        |
|               | Retics ( $\times 10^9/l$ ) |             |                            | 49.3 ± 23.5                |          |                         |                          |                             | 60.0 ± 18.4                        |
|               | Time to DBS (h)            |             |                            | 4.56 ± 6.59                |          |                         |                          |                             | 4.42 ± 7.45                        |

\*Age- and gender-dependent.



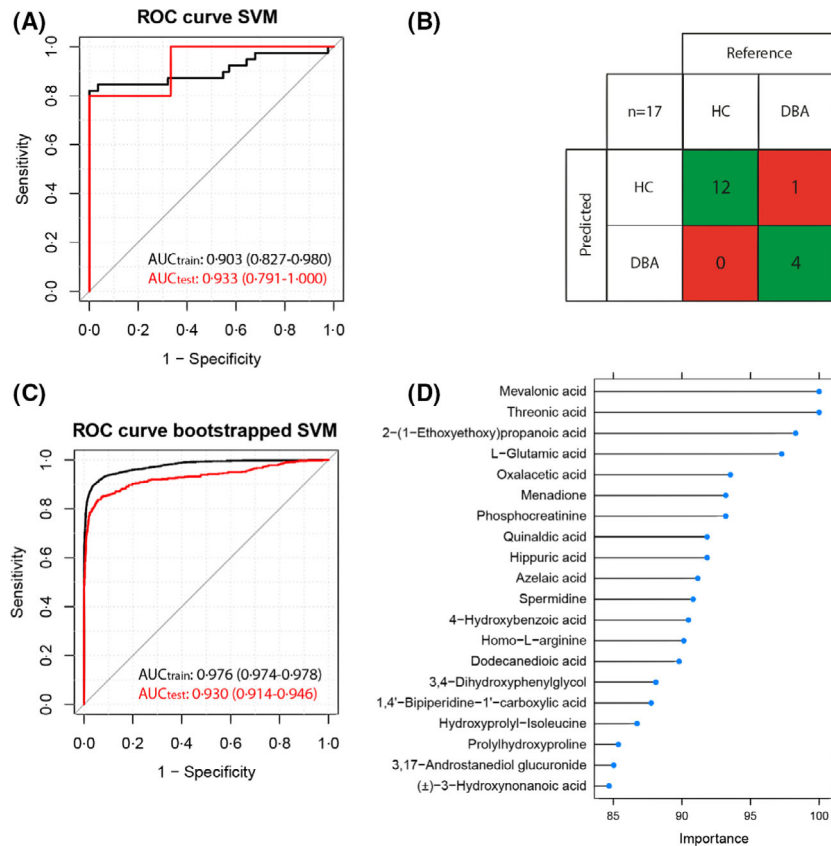
**Fig 1.** Metabolic fingerprint in DBS of DBA patients. (A) Principal component analysis (PCA) of Diamond Blackfan Anaemia (DBA) and healthy controls (HC) displayed with 95% confidence regions. (B) Heatmap of the 25 most significant features identified by *t*-test (raw *P* < 0.0001). The heatmap was created using Euclidian ward-clustering with autoscaling of features. A comprehensive overview of *P*-values and isomers is displayed in Table SI. (C) Partial least squares discriminant analysis (PLS-DA) of DBA and HC displayed with 95% confidence regions. (D) Variable importance in projection (VIP)-scores demonstrating the 25 features contributing the most to separation of patients and controls in the PLS-DA. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

between samples and groups (DBA and HC), initial data exploration was performed by principal component analysis (PCA), an unsupervised technique which does not take the group label into account. This demonstrated equally distributed variance in metabolic profiles, with a certain degree of overlap between DBA and HC (Fig 1A).

Furthermore, using univariate *t*-test analysis, we observed significant differences between groups, including decreased metabolites corresponding in mass to methylmalonylcarnitine (an acyl carnitine), homo-L-arginine, hippuric acid (an acyl glycine) and lipoic acid (an essential cofactor for

mitochondrial enzyme complexes). In addition, we found significantly increased glutamic acid, threonine acid [possibly a metabolite of ascorbic acid (vitamin C)], dodecanedioic acid (a medium-chain fatty acid) and inosine (one of the purine nucleosides) (Fig 1B).

We next performed partial least squares discriminant analysis (PLS-DA), a supervised analysis that takes into account the group label. This resulted in clear separation between DBA and HC, thereby confirming distinct metabolic signatures. (Fig 1C). Metabolites contributing most to the separation are shown in Fig 1D.



**Fig 2.** Machine-learning algorithm predicts DBA based on metabolic profile. (A) Classification performance of samples in training (CV SVM model,  $n = 84$  control, 39 DBA) and test set ( $n = 12$  control, 5 DBA). Note that AUC is a measure of the ability to rank samples according to the probability of class membership, meaning that even falsely classified samples can have a higher rank towards the correct class compared to other samples. (B) Confusion matrix for the prediction of samples from the test set by the SVM model. (C) Classification performance of samples in 100 bootstrapped SVM models (training:  $n = 11\ 901$  control, 5499 DBA; test:  $n = 1381$  control, 519 DBA). (D) Top 20 important features with importance score identified by support vector machine. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

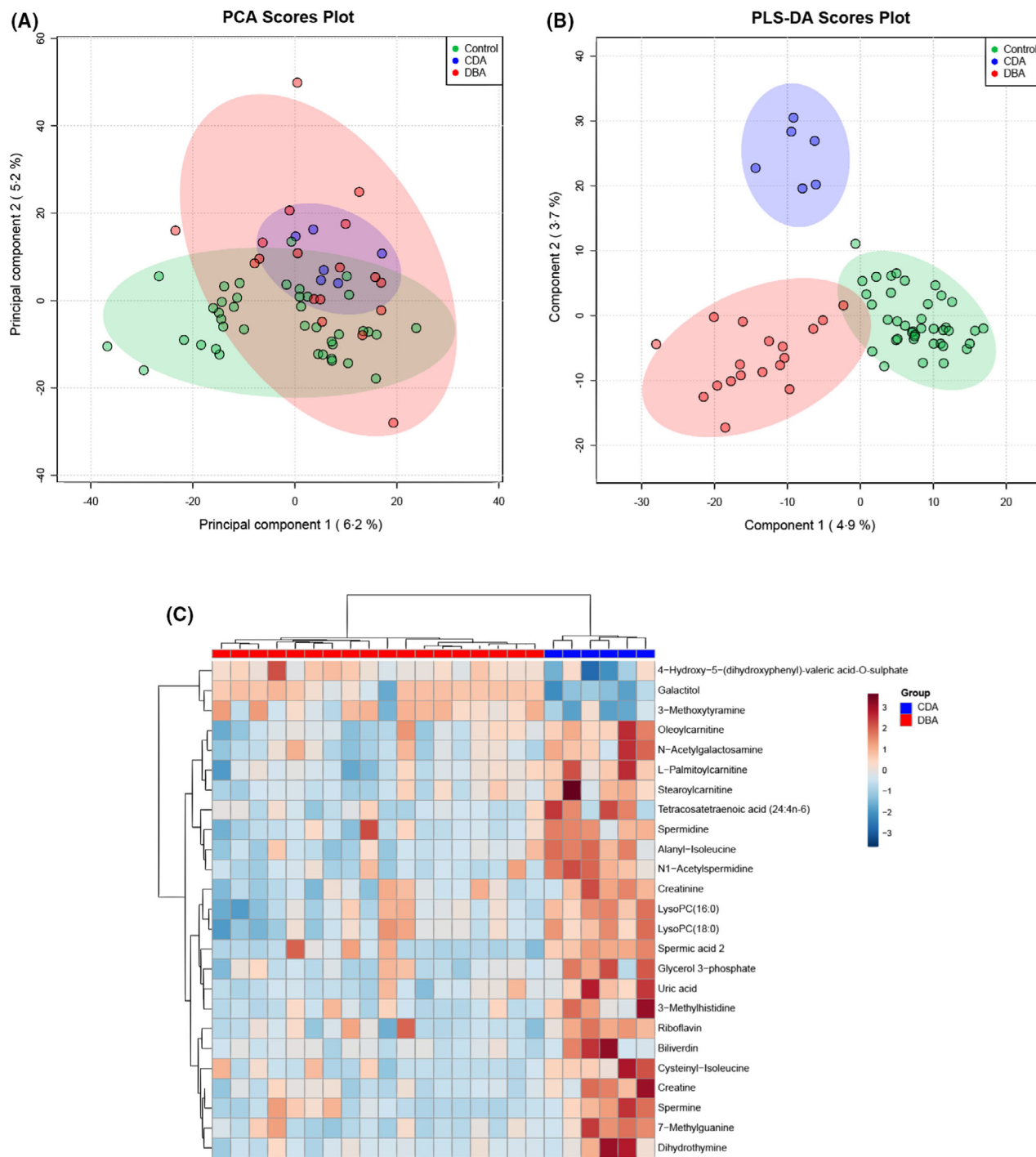
### *A machine-learning algorithm enables prediction of DBA with high accuracy*

To generate a model that can predict the diagnosis of DBA based on the metabolic profile, the dataset was incorporated into a machine-learning algorithm. Using a support vector machine with a linear kernel, a binary classification model was constructed by randomly dividing patients and controls into a 'training' set (28 HC, 13 DBA) or 'test' set (12 HC, 5 DBA), followed by repeated (internal) cross validation (three-fold, three repeats) to identify the optimal hyperplane separating patients from controls. The final model had high performance characteristics with an average accuracy of 91.9%. In addition, receiver operator characteristic curves with the area under the curve were used as performance indicator (Fig 2A, AUC<sub>train</sub>: 0.903 (0.827–0.980)). External model validation was performed by predicting the 'new' control and patient samples from the 'test' set (12 HC, five DBA) (Fig 2A, AUC<sub>test</sub>: 0.933 (0.791–1.000)). For the selected final model, this resulted in the accurate prediction of class for 4/5 patients and all controls (accuracy 94.1%, CI 95%: 71.3–0.99.9) (Fig 2B). To assess the uncertainty of the

final model and its predictive ability, bootstrap resampling (100 repeats) was applied to the entire dataset, resulting in a similarly high prediction performance (AUC<sub>test</sub>: 0.930 (0.914–0.946), and supporting the validity of the presented model (Fig 2C). Important features for classification in this model again included methylmalonylcarnitine, dodecanedioic acid, hippuric acid and glutamic acid (Fig 2D).

### *DBA and CDA patients show a distinct metabolic profile*

To assess disease specificity of the metabolic fingerprint, the DBS metabolome of six patients with congenital dyserythropoietic anaemia (CDA) – a clinically overlapping disorder of erythropoiesis in which proliferation and differentiation of erythroid precursors is affected – was studied (patient characteristics are shown in Table SII). Although the CDA group was relatively small, a distinct metabolic profile was observed for CDA patients compared to both DBA patients and healthy controls, by natural clustering in the PCA plot (Fig 3A) and evident separation of both patient groups in PLS-DA (Fig 3B). The most significant differences between DBA

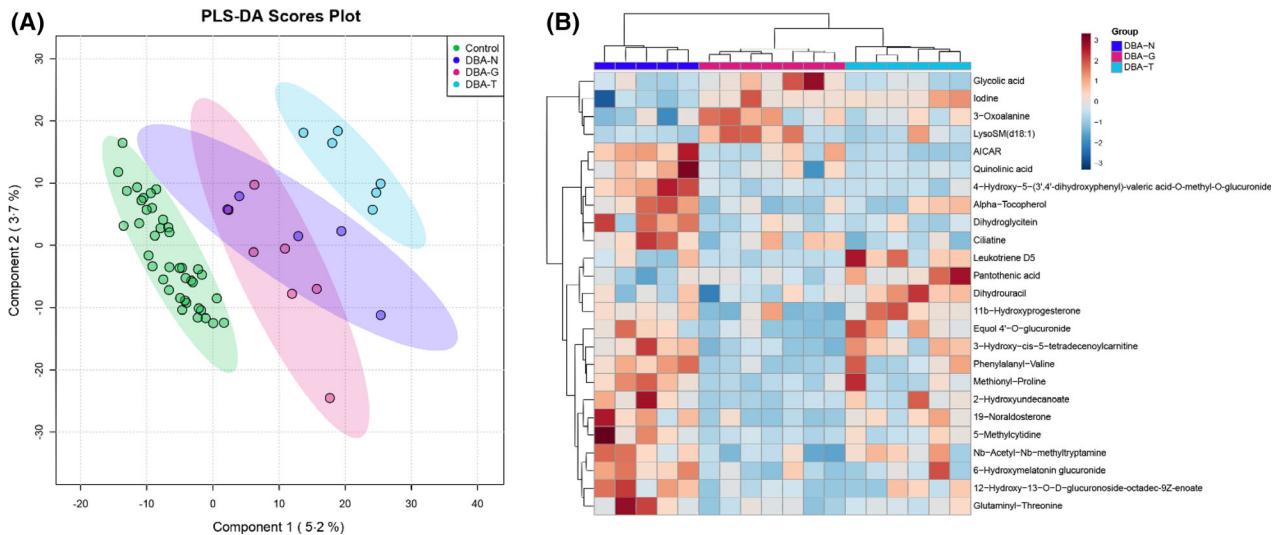


**Fig 3.** Metabolic profile of DBA compared to CDA. (A) PCA plot, and (B) PLS-DA plot for DBA patients compared to CDA patients and healthy controls, displayed with 95% confidence regions. (C) Heatmap of 25 most significant features identified by *t*-test (raw *P* < 0.002) for DBA vs. CDA. The heatmap was created using Euclidian ward-clustering with autoscaling of features. A comprehensive overview of *P*-values and isomers is displayed in Table SI. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and CDA, identified by *t*-test, are shown in a heatmap in Fig 3C, including an increase in acyl carnitines (stearoylcarnitine, oleoylcarnitine and L-palmitoylcarnitine), riboflavin (a vitamin B2 precursor) and polyamines (spermine, spermic acid 2, spermidine and N1-acetylspermidine) in the CDA patients.

*DBS metabolome are instrumental to study differences in therapeutic subgroups*

To investigate the heterogeneity of DBA metabolic profiles in relation to treatment modalities, PCA (data not shown) and



**Fig 4.** Metabolic profile in relation to treatment modality. (A) Partial least squares discriminant analysis (PLS-DA) of non-treated (DBA-N), steroid (DBA-G) and transfused DBA patients (DBA-T) displayed with 95% confidence regions. (B) Heatmap of 25 most significant features identified by t-test (raw  $P$ -value  $<0.02$ ). The heatmap was created using Euclidian ward clustering with autoscaling of features. A comprehensive overview of  $P$ -values and isomers is displayed in Table S1. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

PLS-DA were performed for the entire group of patients and controls. Based on treatment (Table I), patients were divided in non-treated (DBA-N), glucocorticoid-treated- (DBA-G) and transfused groups (DBA-T). In terms of metabolic profile, untreated DBA patients and patients treated with glucocorticoids overlapped and profiles differed the least from controls. Transfused DBA patients were metabolically most distinct from controls (Fig 4A). When controls were removed from the respective analyses, the most significant differences between DBA treatment subgroups, included increased alpha-tocopherol (active vitamin E) in non-treated patients, increased pantothenic acid (vitamin B5) in transfused patients and decreased 3-hydroxy-cis-5-tetradecenoylcarnitine in glucocorticoid-treated patients (Fig 4B).

## Discussion

In this study, we performed untargeted metabolomics in dried blood spots in a relatively large cohort of DBA patients with diverse genotypes and clinical phenotypes. We report here for the first time a specific metabolic fingerprint for DBA.

By incorporating the metabolic fingerprint of DBA in a machine learning algorithm, we demonstrate promising performance characteristics for predicting DBA. This highlights the diagnostic potential of our approach. This is further strengthened by the distinct metabolic profile of DBA and CDA patients, the latter being a disorder with clinical and diagnostic features similar to DBA. Due to the rare nature of CDA and, consequently, the relatively small sample size in this study, it is currently not possible to design a machine learning algorithm that integrates both CDA and DBA patients. The distinct profile in our cohort of CDA patients,

however, supports the hypothesis that the metabolic fingerprint of DBA is disease specific. Further study of a larger number of CDA-samples could enable the construction of a multi-diagnosis model.

Generally, the validity of machine learning-based algorithms is strongly dependent on the input that is used in its construction. Considering that DBA is a rare disease, the current cohort of patients on which our model is based is substantial, yet relatively small and too heterogeneous to generate a powerful prediction model. This is demonstrated by the incorrect classification of one DBA patient with a pathogenic RPS19 mutation yet only mild clinical features. While distinct profiles were identified, and the algorithm performed robustly despite group heterogeneity, increasing sample size (including better age-matched controls) will allow for better stratification of patients. For example, the specific profiles seen for different treatment modalities, illustrate that this approach of DBS metabolomics can be instrumental in investigating subgroups and heterogeneity in DBA. In line with this, the current approach could offer a platform for investigating treatment responses. This is supported by the finding that non-treated patients (those in haematological remission) and glucocorticoid-treated patients cluster most closely to controls, while transfused patients show the most deviant profile. Studying the DBS metabolome in glucocorticoid-treated patients in particular could identify determinants of steroid response, and help predict steroid response in patients. To this end, future investigations should include sampling patients before the start of – and during treatment with – glucocorticoids.

Targeting cellular metabolism is increasingly and successfully being explored as a therapeutic option for various forms of hereditary anaemias, and metabolic insights have proven

vital in these developments.<sup>15–18</sup> In particular, the amino acid leucine is currently being explored as a potential modulator of protein synthesis in DBA patients.<sup>19</sup> Recently, results of the first clinical trial have been published, showing a partial or complete haematological response in 7/43 patients (16.3%), and more strikingly, positive effects on growth in 9/25 (height) and 11/25 (weight) patients, respectively.<sup>20</sup> How leucine exerts its effect on both erythropoiesis and growth remains to be defined, but it has been assumed that it involves upregulating of ribosome biosynthesis via the mTOR pathway, as well as improved translational efficiency through activation of translation-initiation factors.<sup>21</sup> While leucine levels in our analyses did not differ significantly between DBA patients and healthy controls (data not shown), increasing insights into the metabolic disturbances that occur in DBA could lead to the identification of new therapeutic targets.

Interestingly, inosine – a purine nucleoside – emerged in the DBA fingerprint among the top metabolites significantly differing between patients and controls. Although the long-recognised increased activity of erythrocyte adenosine deaminase – an enzyme of the purine salvage pathway – has remained elusive to date in the majority of DBA patients, our findings suggest that purine metabolism might be more broadly involved in DBA disease biology. In addition, the finding of increased alpha-tocopherol in the non-treated patients (i.e. patients who are recognised as being in ‘haematological remission’) is interesting, as vitamin E facilitates intracellular scavenging of ROS, and its levels in blood are known to reflect redox status.<sup>22,23</sup> Hence, a better anti-oxidant status may be associated with the haematological remission phenotype in these patients.

Future studies are required to determine and validate these metabolic disturbances at a (red) cellular level as opposed to the whole blood metabolome that we study here.

In conclusion, we report on a new application of untargeted metabolomics in dried blood spots, a minimally invasive approach with diagnostic potential in Diamond Blackfan Anaemia, which could be instrumental in investigating clinical phenotypes and treatment response.

## Acknowledgements

The authors thank Fini de Gruyter for her technical support in Bio-informatics. This study was supported in part by research funding from Metakids (Grant No. 2017-075) to Judith Jans.

## Author contributions

BD and MBr contributed to collection, analysis and interpretation of the data, and wrote the first manuscript. WS, EN and NV were all involved in the study design and carefully revised the manuscript. EB was actively involved in collecting patient samples and carefully revised the manuscript. MBa, JJ

and RW were principal investigators and were involved in all aspects of the study, including design, collection and interpretation of the data, as well as revising and co-writing the manuscript.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table SI.** Overview of *P*-values and isomers.

**Table SII.** Patient characteristics of CDA patients.

## References

- Diamond L, Blackfan K. Hypoplastic anemia. *Am J Dis Child.* 1938;**56**:464–7.
- Dianzani I, Garelli E, Ramenghi U. Diamond-Blackfan anemia: a congenital defect in erythropoiesis. *Haematologica.* 1996;**81**(6):560–72.
- Bartels M, Bierings M. How I manage children with Diamond-Blackfan anaemia. *Br J Haematol.* 2019;**184**(2):123–33.
- Narla A, Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. *Blood.* 2010;**115**(16):3196–205.
- Ulirsch JC, Verboon JM, Kazerounian S, Guo MH, Yuan D, Ludwig LS, et al. The genetic landscape of Diamond-Blackfan Anemia. *Am J Hum Genet.* 2018;**103**(6):930–47.
- Glader BE, Backer K, Diamond LK. Elevated erythrocyte adenosine deaminase activity in congenital hypoplastic anemia. *New Engl J Med.* 1983;**309**(24):1486–90.
- Smetanina NS, Mersyanova IV, Kurnikova MA, Ovsyannikova GS, Hachatrian LA, Bobrynina VO, et al. Clinical and genomic heterogeneity of Diamond Blackfan anemia in the Russian Federation. *Pediatr Blood Cancer.* 2015;**62**(9):1597–600.
- van Dooijeweert B, van Ommen CH, Smiers FJ, Tamminga RY, te Loo MW, Donker AE, et al. Pediatric Diamond-Blackfan anemia in the Netherlands: an overview of clinical characteristics and underlying molecular defects. *Eur J Haematol.* 2018;**100**(2):163–70.
- Vlachos A, Ball S, Dahl N, Alter BP, Sheth S, Ramenghi U, et al. Diagnosing and treating Diamond Blackfan anaemia: results of an International Clinical Consensus Conference. *Br J Haematol.* 2008;**142**(6):859–76.
- Patti GJ, Yanes O, Siuzdak G. Innovation: metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol.* 2012;**13**(4):263–9.
- de Sain-van der Velden MGM, van der Ham M, Gerrits J, Prinsen HCMT, Willemsen M, Pras-Raves ML, et al. Quantification of metabolites in dried blood spots by direct infusion high resolution mass spectrometry. *Anal Chim Acta.* 2017;**979**:45–50.
- Haijes H, Willemsen M, van der Ham M, Gerrits J, Pras-Raves M, Prinsen H, et al. Direct infusion based metabolomics identifies metabolic disease in patients' dried blood spots and plasma. *Metabolites.* 2019;**9**(1):12.
- Van Dooijeweert B, Broeks MH, Verhoeven-Duif NM, Van Beers EJ, Nieuwenhuis EES, Van Solinge WW, et al. Untargeted metabolic profiling in dried blood spots identifies disease fingerprint for pyruvate kinase deficiency. *Haematologica.* 2020. <https://doi.org/10.3324/haematol.2020.266957>
- Chong J, Wishart DS, Xia J. Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Curr Protoc Bioinformatics.* 2019;**68**(1):e86.
- Cox SE, Ellins EA, Marealle AI, Newton CR, Soka D, Sasi P, et al. Ready-to-use food supplement, with or without arginine and citrulline, with daily chloroquine in Tanzanian children with sickle-cell disease: a double-blind, random order crossover trial. *Lancet Haematol.* 2018;**5**(4):e147–e160.
- Darghouth D, Koehl B, Madalinski G, Heilier JF, Bovee P, Xu Y, et al. Pathophysiology of sickle cell disease is mirrored by the red blood cell metabolome. *Blood.* 2011;**117**(6):e57–e66.



17. Morris CR, Kuypers FA, Lavrisha L, Ansari M, Sweeters N, Stewart M, et al. A randomized, placebo-controlled trial of arginine therapy for the treatment of children with sickle cell disease hospitalized with vaso-occlusive pain episodes. *Haematologica*. 2013;**98**(9):1375–82.
18. Grace RF, Rose C, Layton DM, Galactéros F, Barcellini W, Morton DH, et al. Safety and efficacy of mitapivat in pyruvate kinase deficiency. *New Engl J Med*. 2019;**381**(10):933–44.
19. Pospisilova D, Cmejlova J, Hak J, Adam T, Cmejla R. Successful treatment of a Diamond-Blackfan anemia patient with amino acid leucine. *Haematologica*. 2007;**92**(5):e66–e67.
20. Vlachos A, Atsidaftos E, Lababidi ML, Muir E, Rogers ZR, Alhushki W, et al. L-leucine improves anemia and growth in patients with transfusion-dependent Diamond-Blackfan anemia: results from a multicenter pilot phase I/II study from the Diamond-Blackfan Anemia Registry. *Pediatr Blood Cancer*. 2020;**67**(12):e28748.
21. Xu B, Gogol M, Gaudenz K, Gerton JL. Improved transcription and translation with L-leucine stimulation of mTORC1 in Roberts syndrome. *BMC Genom*. 2016;**17**:25.
22. Margaritelis NV, Veskoukis AS, Paschalis V, Vrabas IS, Dipla K, Zafeiridis A, et al. Blood reflects tissue oxidative stress: a systematic review. *Biomarkers*. 2015;**20**(2):97–108.
23. Niki E. Role of vitamin E as a lipid-soluble peroxy radical scavenger: in vitro and in vivo evidence. *Free Radic Biol Med*. 2014;**66**:3–12.