Showcasing research from Professor Stefano Materazzi, Department of Chemistry – "Sapienza" University of Rome, P.le A.Moro 5, 00185 Rome, Italy

Differential diagnosis of hereditary hemolytic anemias in a single multiscreening test by TGA/chemometrics

A multi-screening test based on the coupling of thermogravimetry and chemometrics was optimized for the differential diagnosis of hereditary hemolytic anemias.

As featured in:

See Stefano Materazzi *et al.*, *Chem. Commun.*, 2020, **56**, 7557.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

Registered charity number: 207890

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2020, **56**, 7557

Received 24th April 2020,
Accepted 4th June 2020

DOI: 10.1039/d0cc02948c

rsc.li/chemcomm

Differential diagnosis of hereditary hemolytic anemias in a single multiscreening test by TGA/chemometrics†

Roberta Risoluti,[a] Patrizia Caprari,[b] Giuseppina Gullifa,[a] Francesco Sorrentino,[c] Laura Maffei,[c] Sara Massimi,[b] Elena Carcassi[a] and Stefano Materazzi *[a]

A multi-screening test based on the coupling of thermogravimetry and chemometrics was optimized for the differential diagnosis of hereditary hemolytic anemias. The novel test is able to simultaneously perform a simple and fast diagnosis of sickle cell anemia, thalassemia, hereditary spherocytosis and hereditary elliptocytosis in a single analysis of a few microliters of non-pretreated whole blood. The thermogravimetric profile of blood from patients affected by such disorders was found to be characteristic of a specific anemic status or a disorder due to membrane defects. In addition, chemometric tools were used to validate a model of prediction to process the thermogravimetric curves and to obtain in 1 hour an accurate diagnosis. The effectiveness of the novel test was evaluated by comparing results with the confirmatory analyses specific for each disorder. The TGA/chemometric test made it possible to perform a first level test of congenital erythrocyte defects, including the hemoglobinopathies and disorders due to membrane defects with the same accuracy of confirmatory analyses obtained by molecular investigation. In addition, the novel test was used for the diagnosis of a number of Italian difficult cases, including neonatal patients for which the conventional screening tests did not manage to obtain a diagnosis confirming the high prediction ability of the single multiscreening test.

Hereditary hemolytic anemias are a group of heterogeneous diseases caused by abnormalities intrinsic to the erythrocytes, including defects in the structure of the red cell membrane, as in hereditary spherocytosis and elliptocytosis,[1] or the presence of abnormal hemoglobins or impaired synthesis of globin chains, as in the hemoglobinopathy syndromes, such as sickle cell anemias and thalassemia.[2] The features of these congenital anemias include decreased red blood cell counts and hemoglobin content, increased reticolocyte count, increased unconjugated bilirubin and lactate dehydrogenase activity, and reduced haptoglobin. Although the alterations of RBC morphology on the peripheral blood smear can be observed, these laboratory features are not able to characterize the disease, since they are common to all hereditary hemolytic diseases.[3] As a consequence, differential diagnosis can be difficult and is generally carried out by applying different diagnostic protocols, specific for each type of congenital erythrocyte defect.[4] The protocol for the diagnosis of hereditary spherocytosis and elliptocytosis provides the red cell morphological examination, the red cell indices, the osmotic fragility, and flow cytometry, as first level methods, and the erythrocyte membrane protein electrophoresis, and the genomic analysis to characterize the various red cell membrane defects.[5]

The identification of patients suffering from hemoglobinopathies, such as sickle cell anemias and thalassemia syndromes requires a diagnostic protocol divided into two levels: the 1st level tests consist of the whole blood count test, the assessment of the martial state and the study of hemoglobin fractions in HPLC. These tests are mainly aimed at identifying the healthy carrier state but, in some cases, make is possible to identify the disease and to define the presence of a hemoglobin variant.[6] Subsequent molecular investigations (2nd level analysis) allow defining the genetic defect and correlating the clinical phenotype with the specific genotype. As a result, the correct diagnosis may require time and costs with a consequent delay in the administration of the correct (personalized) therapy.

In this study, a novel multiscreening test able to simultaneously perform the diagnosis of hereditary spherocytosis, hereditary elliptocytosis, sickle cell anemia and thalassemia in a single analysis, is proposed. In particular, the test uses thermogravimetry associated with chemometrics on a few microliters of whole blood without requiring any sample pretreatment.

[a] *Department of Chemistry – "Sapienza" University of Rome, P.le A.Moro 5, 00185 Rome, Italy. E-mail: stefano.materazzi@uniroma1.it; Fax: +39-0649387137; Tel: +39-0649913616*

[b] *National Centre for the Control and Evaluation of Medicines, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy*

[c] *Thalassemia Unit, S. Eugenio Hospital, Piazzale dell'Umanesimo 10, 00144 Rome, Italy*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cc02948c

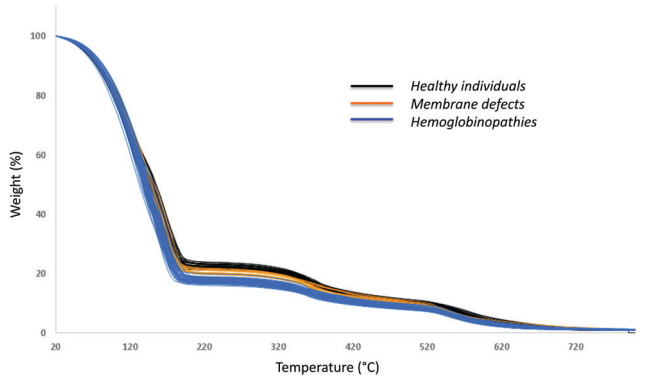

Fig. 1 Thermogravimetric curves of healthy individuals (black), patients affected by membrane defects (orange) and hemoglobinopathies (SCA and thalassemia patients, light blue).

Hematological characterization: patients were classified by evaluating the characteristic features for hemolytic anemias due to hemoglobinopathies and membrane defects. Table S1 (ESI†) shows the hematological data expressed as mean ± standard deviation (SD) while the comparison among groups was performed by the analysis of variance (ANOVA) followed by the Tukey test to assess significances among classes. A detailed description of patients and the discussion of the haematological characterization are reported in the ESI.†

TGA/Chemometrics multiscreening test: blood samples from all the involved subjects were processed by thermogravimetric analysis in order to evaluate the TG profile of blood samples when heated under controlled conditions. The characteristic thermal behaviour of patients affected by hemolytic anemias was estimated and compared to the thermal profile of healthy subjects, as reported in Fig. 1.

The analysis of the TG curves showed a different thermally induced decomposition of blood, especially according to the first releasing step. In particular, the percentage weight losses as a function of the temperature were calculated as the mean of the population and standard deviation and reported in Table S2 (ESI†). Significances were calculated by the analysis of the variance (ANOVA) and the Tukey test was used to compare the classes. Although the groups of patients affected by anemias exhibited higher standard deviations, significant differences (sample mean difference among classes exceeding the Tukey critical value) may be observed in the acquired TG curves under oxidative conditions. In particular:

– Thalassemia and SCA patients presented the highest values of water content with a consequent decrease in the corpuscular fraction, shown from the second and the third weight losses. This result is in accordance with the characteristic features of the two populations belonging to the same class of hemoglobinopathies; both the classes were found to be statistically different from the others;

– The total amount of the water in the group of HS/HE patients is lower than the hemoglobinopathies and was found to be close to the healthy subjects.

– The water content percentage was found to be significantly different between healthy and hemoglobinopathy subjects, with group mean differences exceeding the Tukey value both for thalassemic subjects and SCA ones. In addition, no significant differences in water content were observed between thalassemic and SCA.

– The percentage of the bulk water in patients affected by hereditary disorders is higher than the healthy subjects with consequent higher values of the bulk/bound water ratio that was found to be not significant. In contrast, the bulk/bound water ratio significantly differs between thalassemic and SCA groups.

Data reported in Table S2 (ESI†) also suggested that not all the considered thermogravimetric processes contributed significantly in differentiating subjects. To this aim, a chemometric investigation of the most descriptive variables considering the entire range of temperature was performed by chemometric analysis of the thermogravimetric curves. In fact, by combining the thermogravimetric analysis with chemometric methods of multivariate statistical analysis such as Principal Component Analysis (PCA), it is possible to study simultaneously the correlations between the thermogravimetric profiles of different blood samples. The resulting scores plot is reported in Fig. 2, where symbols represent each sample and colors are used to indicate the disease. The chemometric analysis of the TG curves, after mean centering correction, made it possible to distinguish the healthy subjects (black squares) from patients affected by an anemic status by considering the results of the scores of PC1 × PC2 (overall explained variance of 95%). In addition, samples were found to be further differentiated in membrane defects and hemoglobinopathies respectively reported as orange circles and light blue triangles. The chemometric investigation of the factor loadings also suggested that both the water content and the bulk/bound water ratio may be selected to discriminate patients: in fact, with respect to previous reported data,[7] the water content is not sufficient enough to differentiate thalassemia from SCA, requiring the calculation of all the TG processes in the entire range of temperatures.

Such a result suggested the possibility to optimize a novel test model for the simultaneous identification of the anemic status and the consequent prediction of the hereditary disease for a differential diagnosis.

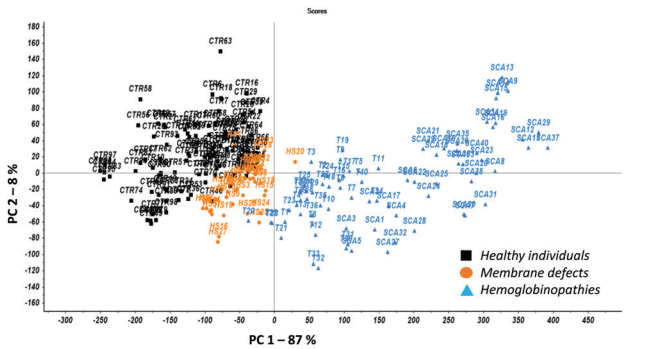

Fig. 2 Resulting scores plot of healthy individuals (black), and subjects affected by membrane defects (orange) and hemoglobinopathies (SCA and thalassemia patients, light blue).

Differential multiscreening test: an innovative automated multiscreening test for a multilevel diagnosis was developed and validated by the mean of chemometric tools. In particular the Partial Least Square Discriminant Analysis (PLS-DA) was involved as a predictive model after PCA investigation. With the aim of proposing a test as accurate as possible for different hereditary red blood diseases, all the collected data were further divided into four groups: the healthy subjects, patients affected by a membrane defect (including HS and HE), patients affected by SCA and thalassemia. Preliminarily, a PCA was calculated to evaluate the sample location in the plot and the possibility to differentiate SCA patients from the thalassemic ones. The resulting scores plot is reported in Fig. 3, where the mean centering pretreatment was used to compare data after variable selection. The loading plots of PC1 and PC2 are reported in Fig. S1 (ESI†).

To perform the PLS-DA model, the entire dataset was split into the training set and evaluation set by the cross validation method.[8] This procedure will ensure that each sample is used to estimate the prediction ability of the multiscreening model. In order to increase separation of samples, and thus to improve sensitivity, specificity and efficiency of the test, a number of TG curve pretreatments were investigated and the effectiveness of each pretreatment was evaluated. The model's performances were provided by calculating the Sensitivity (True Positive rate, TP), Specificity (True Negative, TN) and Efficiency Rate.[9] In particular, the effect of mean centering and SNV was tested and also the improvement increased by the calculation of the first derivative of the TG curve prior to mean centering and SNV, respectively. The calculation of the figures of merit is reported in Fig. S2 and Tables S3–S5 (ESI†), that provided suitable results in calibration and cross validation for all the investigated pre-treatments. In particular, the mean centering made it possible to obtain suitable values of efficiency (not less than 86%) and the required values of sensitivity and specificity (not lower than 80% in prediction). Therefore, although in some cases the involvement of the first derivative on the mean centering pretreatment provided the highest values of sensitivity in calibration and cross-validation, an unacceptable specificity was achieved in the prediction of the external dataset. The resulting 3D plot of the PLS-DA model is reported in Fig. 4, while the figures of merits, calculated by the definitive method using 7 latent variables, provided a sensitivity of about 80% for HS/HE and 100% for the remaining classes. The specificity was about 83% for the healthy subjects and 100% for HS/HE, thalassemic and SCA patients. Finally, the model ensured the correct differential diagnosis may be performed with an efficiency rate of about 96.8% for healthy subjects, 89.4% for HS/HE subjects and 100% for both thalassemic and SCA patients.

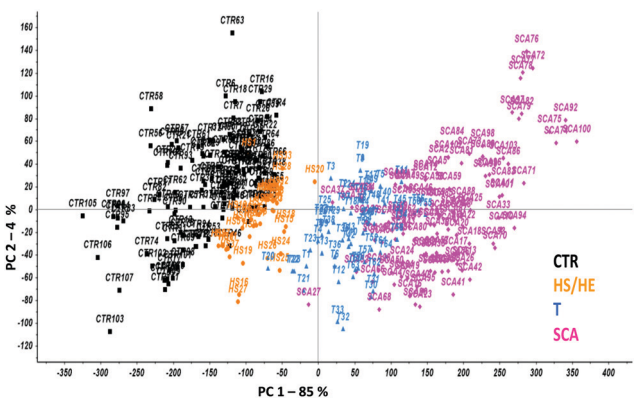

Fig. 3 Scores plot from PCA for healthy (black), membrane defects (orange), thalassemic (light blue) and SCA (pink) patients.

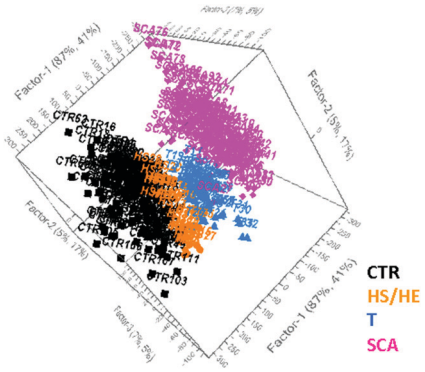

Fig. 4 3-D plots of the PLS-DA regression model for healthy (black), membrane defect (orange), thalassemic (light blue) and SCA (pink) patients.

Diagnosis of rare hereditary disorders: the hereditary hemolytic anemias generally present in the newborn period and in early childhood, when it may be difficult to perform differential diagnosis, but patients affected by mild forms may show clinical signs only in adulthood. The multiscreening test was used to investigate some cases of anemias of difficult diagnosis (some of them were children), characterized by hyperbilirubinemia, negative Coombs tests, and increased LDH value that suggested the presence of an erythrocyte congenital defect as the cause of the hemolytic anemia. Fourteen patients with hemolytic anemia were analyzed by TGA and chemometric analysis after the screening methods of congenital hemolytic anemias from hemoglobin disorders, and defects of membrane proteins had not found the cause of anemia. The TGA/chemometrics multiscreening test showed the presence of two cases of HS, five patients with thalassemia, and seven patients with sickle cell anemia. The TGA screening results were confirmed by the 2nd level analysis of the membrane proteins and molecular analysis of the globin genes. All the analyzed patients were found to be correctly predicted by the multiscreening test that permitted a differential diagnosis to point out a specific therapeutic treatment.

Patient enrollment and characterization: a number of 324 subjects were considered in this study: 120 healthy donors (with mean age ± standard deviation of 40 ± 11 years) were characterized at the National Health Institute of Rome, while 204 patients affected by hemolytic disorders were followed through diagnosis, management and therapies at the Thalassemia Unit of S. Eugenio Hospital in Rome. All experiments were performed in compliance with the guidelines established by the Ethical Committee for human subject studies (Helsinki Declaration of 1975, revised in 2008)

and the study was approved by Comitato Etico Roma 2 of the S. Eugenio Hospital of Rome. Written informed consent was collected from patients for experimentation. Characterization of patients was performed by using a hematological analyzer and thermogravimetric analysis (TGA): the description of patients and the analytical conditions are reported in the ESI.†

Analytical workflow and chemometrics: the novelty of this test consists of the possibility to rapidly perform a differential diagnosis in a single analysis, in one hour of analysis time and with the accuracy of the second level test. In addition, positivity of this test will rapidly address patients to the specific confirmatory analyses and make it possible to provide the correct therapeutic protocol.

Chemometric tools were used to compare data and to develop a novel test able to perform a differential diagnosis of hereditary hemolytic anemia in 30 μl of whole blood. In particular, Principal Component Analysis (PCA)[10] was used as an exploratory method, while Partial Least Square Linear Discriminant Analysis (PLS-DA)[11] was considered as the regression model of prediction. Collected data were subjected to a random splitting process and the cross validation method was performed. A chemometric study was planned with the aim of evaluating the best performing model to predict the disorders. In particular, the model's performances were assessed by calculating the False Positive Rate (FP), False Negative Rate (FN), Sensitivity (True Positive rate, TP), Specificity (True Negative, TN) and Efficiency Rate, as defined in ref. 12. The entire dataset of samples was divided into training set and evaluation set by means of cross validation and the identification of outliers was performed by using the non-linear iterative partial least squares (NIPALS) algorithm and by considering the Mahalanobis distance. The sensitivity and false negative rate were assessed by cross-validation procedures in order to ensure independency, while the specificity and the false positive rate were calculated by predicting subjects affected by a different disorder. Diagnostics and acquisition of the thermogravimetric data were carried out by Pyris software (Thermo Fisher Scientific Inc., Waltham, MA, USA) as ASCII files, which were then imported into the Unscrambler package to perform statistical analysis.

In this work, a novel multiscreening test is proposed and validated to perform a differential diagnosis of hereditary red blood cells disorders. In particular, the model was demonstrated to be able to identify the congenital anemic status due to a defect as in hereditary spherocytosis and elliptocytosis, or to the presence of abnormal hemoglobins or impaired synthesis of globin chains, as in the hemoglobinopathies and thalassemia syndromes.

Positivity of this test will rapidly lead patients to the specific confirmatory analyses and make it possible to provide the correct therapeutic protocol, after 1 hour and using only a few microliters of non pre-treated whole blood.

In addition, the test was also able to perform the prediction of difficult cases of congenital defects that the conventional first level tests did not identify, opening new frontiers to an accurate differential diagnosis at the first level test especially for newborn screening programs.

This manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Conflicts of interest

The authors declare no competing financial interest.

Notes and references

1 L. Da Costa, J. Galimand, O. Fenneteau and N. Mohandas, *Blood Rev.*, 2013, **27**(4), 167–178; S. Rocha, E. Costa, P. Rocha-Pereira, F. Ferreira, E. Cleto, J. Barbot, A. Quintanilhabh, L. Belo and A. Santos-Silva, *Clin. Biochem.*, 2011, **44**(13), 1137–1143; S. Perrotta, P. G. Gallagher and N. Mohandas, *Lancet*, 2008, **372**, 1411–1426.

2 S. Azar and T. E. Wong, *Med. Clin. North Am.*, 2017, **101**, 375–393; D. C. Rees, T. N. Williams and M. T. Gladwin, *Lancet*, 2010, **376**, 2018–2031; G. J. Kato, F. B. Piel, C. D. Reid, M. H. Gaston, K. Ohene-Frempong, L. Krishnamurti, W. R. Smith, J. A. Panepinto, D. J. Weatherall, F. F. Costa and E. P. Vichinsky, *Nat. Rev.*, 2018, **4**, 1–22; D. J. Weatherall and J. B. Clegg, *The Thalassemia Syndromes*, Oxford, Blackwell Science, UK, 2001. A. Cao, R. Galanello, *Int. J. Gen. Med.*, 2010, 12, 61–76; S. L. Thein, *Cold Spring Harbor Perspect. Med.*, 2013, **3**(5), a011700.

3 A. N. Schechter, *Blood*, 2008, **112**, 3927–3938; E. Kohne, *Dtsch. Arztebl. Int.*, 2011, **108**, 532–540.

4 B. Giardine, J. Borg, E. Viennas and C. Pavlidic, *Nucleic Acids Res.*, 2014, **42**, 1063–1069; A. Mosca, R. Palearo, D. Leone and G. Ivaldi, *Clin. Biochem.*, 2009, **42**(18), 1797–1801.

5 S. Rocha, E. Costa, P. Rocha-Pereira, F. Ferreira, E. Cleto, J. Barbot, A. Quintanilhabh, L. Belo and A. Santos-Silva, *Blood Cells, Mol., Dis.*, 2011, **46**(15), 166–170; K. Haley, *Med. Clin. North Am.*, 2017, **101**, 361–374; M. J. King and A. Zanella, *Int. J. Lab. Hematol.*, 2013, **35**(3), 237–243.

6 B. H. Lubin, H. E. Witkowska and K. Kleman, *Clin. Biochem.*, 1991, **24**(4), 363–374; D. N. Greene, C. P. Vaughn, B. O. Crews and A. M. Agarwal, *Clin. Chim. Acta*, 2015, **439**, 50–57; A. Dasgupta and A. Wahed, *Certification and Clinical Practice*, 2014, 363–390.

7 R. Risoluti, M. Materazzi, F. Sorrentino, L. Maffei and P. Caprari, *Talanta*, 2016, **159**, 425–432; R. Risoluti, S. Materazzi, F. Sorrentino, C. Bozzi and P. Caprari, *Talanta*, 2018, **183**, 216–222.

8 L. Xu, O. Hu, Y. Guo, M. Zhang, D. Lu, C. Cai, S. Xie, M. Goodarzi, H. F. Fu and Y. B. She, *Chemom. Intell. Lab. Syst.*, 2018, **183**, 29–35; M. Forina, P. Oliveri and M. Casale, *Chemom. Intell. Lab. Syst.*, 2010, **102**, 110–122; R. G. Brereton, *Chemometrics for Pattern Recognition*, Wiley, 2009.

9 P. Oliveri and G. Downey, *TrAC, Trends Anal. Chem.*, 2012, **35**, 74–86; M. Otto, *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, Wiley-VCH, Weinheim, 3rd edn, 2017; K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, Cambridge, Mass, 2012; S. Materazzi, G. Gullifa, M. A. Fabiano, P. Frati, A. Santurro, M. Scopetti, V. Fineschi and R. Risoluti, *J. Therm. Anal. Calorim.*, 2017, **130**(1), 549–557.

10 D. L. Massart, B. G. M. Vandeginst, L. M. C. Buydens, S. De Jong, P. L. Lewi and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics. Part B*, 1998, vol. 20B, pp. 88–103; M. Ferreiro-González, E. Espada-Bellido, L. Guillén-Cueto, M. Palma, C. M. Barroso and G. F. Barbero, *Talanta*, 2018, **188**, 288–292; R. Risoluti, A. Gregori, S. Schiavone and S. Materazzi, *Anal. Chem.*, 2018, **90**(7), 4288–4292.

11 A. Savitzky and M. J. E. Golay, Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.*, 1964, **36**, 1627–1639; M. W. Barker, *J. Chemom.*, 2003, **17**, 166–173; D. L. Massart, B. G. M. Vandeginst, L. M. C. Buydens, S. De Jong, P. L. Lewi and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics. Part B*, 1998, **vol. 20B**, 213–220.

12 M. I. Lopez, M. P. Callao and I. Ruisanchez, *Anal. Chim. Acta*, 2015, **891**, 62–72; M. P. Callao and I. Ruisanchez, *Food Control*, 2018, **86**, 283–293.

