Electronic Supplementary Material (ESI) for Analytical Methods.

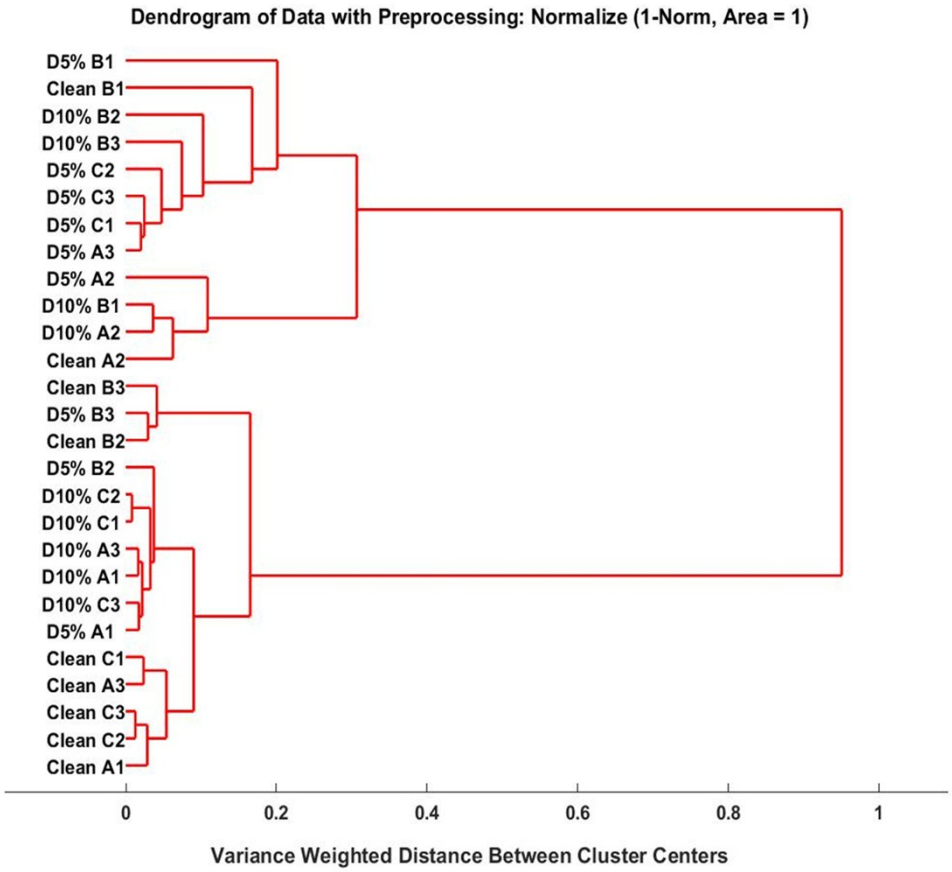
Supporting Information for

Informatics Analysis of Capillary Electropherograms of Autologously Doped and Undoped Blood

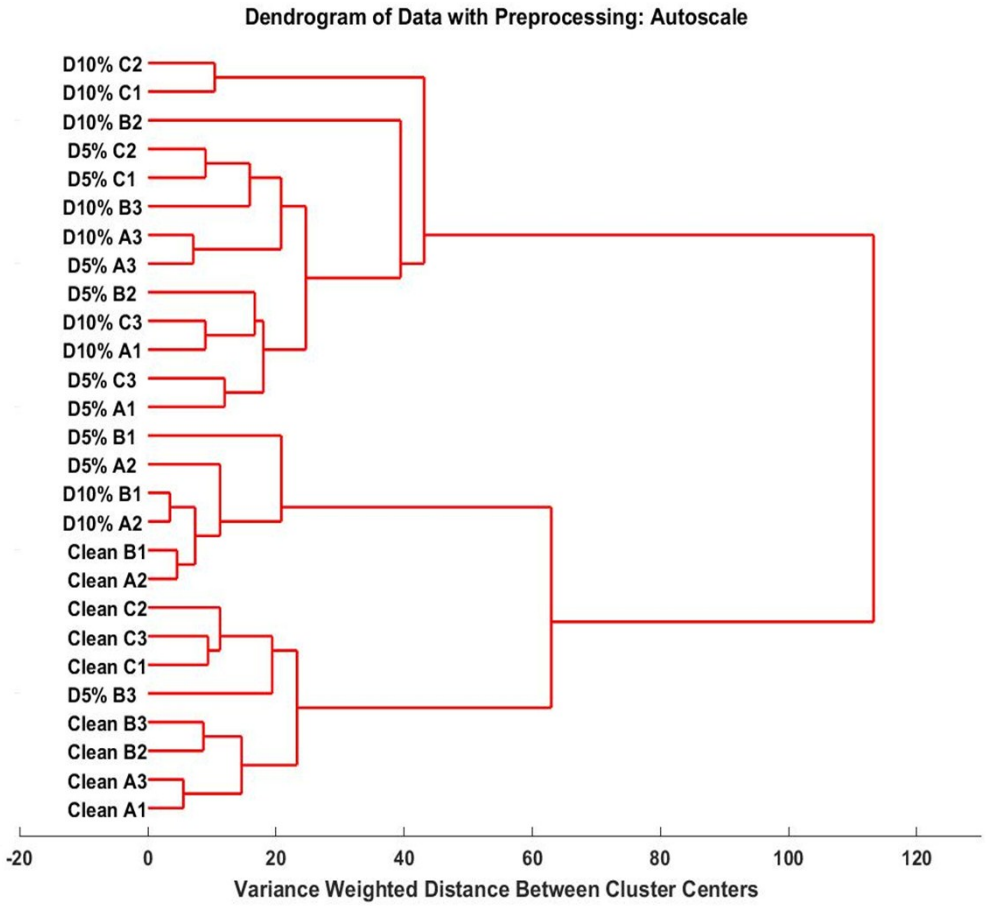
Shiladitya Chatterjee[1], Sean C. Chapman[1], George H. Major, [1] Denis L. Eggett[1], Barry M. Lunt[1], Christopher R. Harrison[2], Matthew R. Linford[1]

[1] Brigham Young University, Provo, Utah 84602, USA

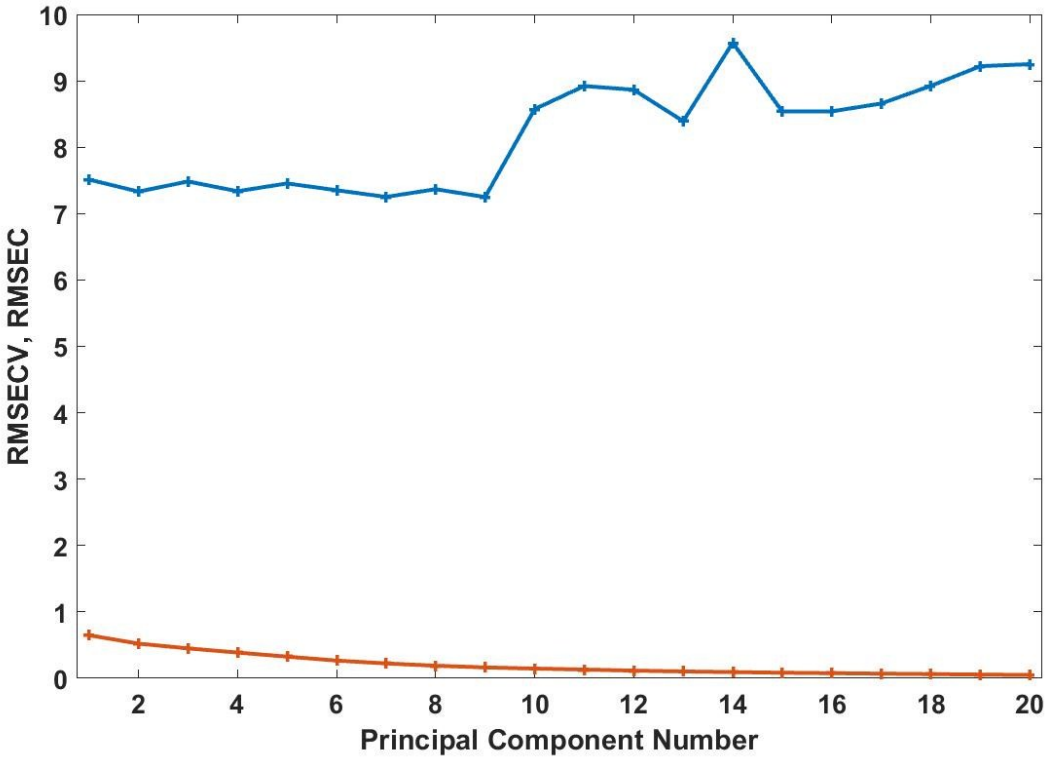
[2]San Diego State University, San Diego, CA 92182-1030, USA



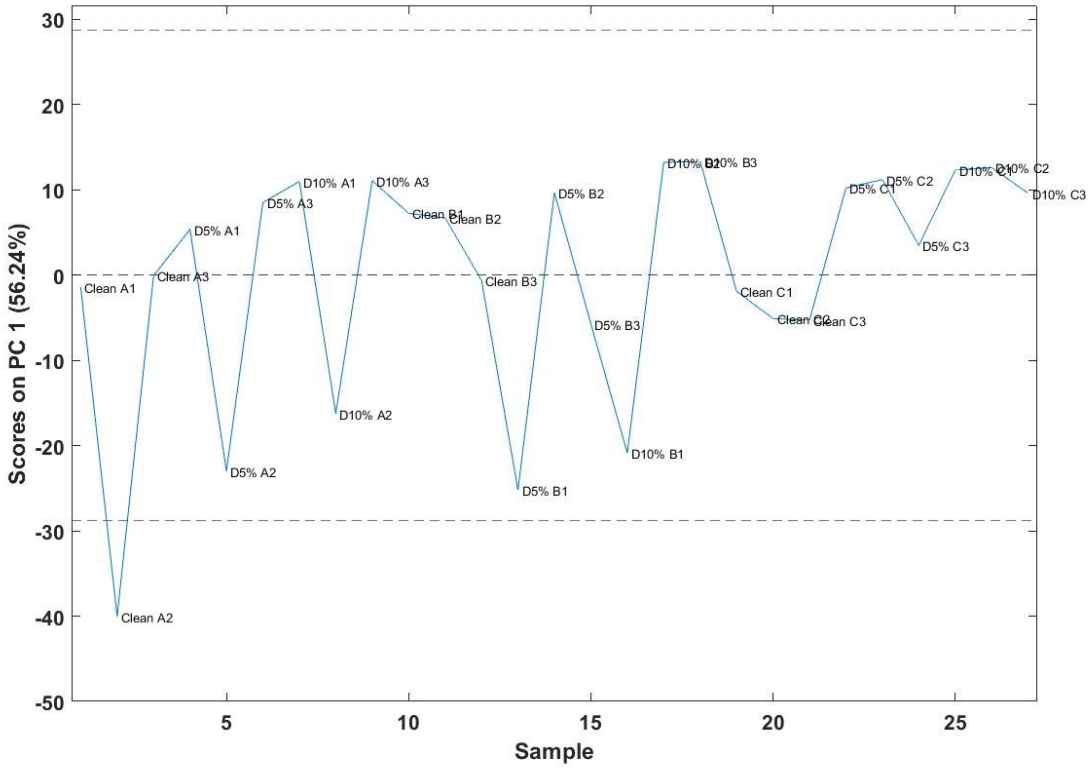
Supporting Information Figure 1. Dendrogram from a cluster analysis of the 0, 5, and 10 % electropherograms under consideration in this study. The data were pre-processed using range-selection followed by normalization (1-norm). 'Clean', 'D5%', and 'D10%' represent the 0, 5, and 10 % samples, and 'A', 'B', and 'C' represent the three subjects. The numbers after 'A', 'B', or 'C' represent replica runs.



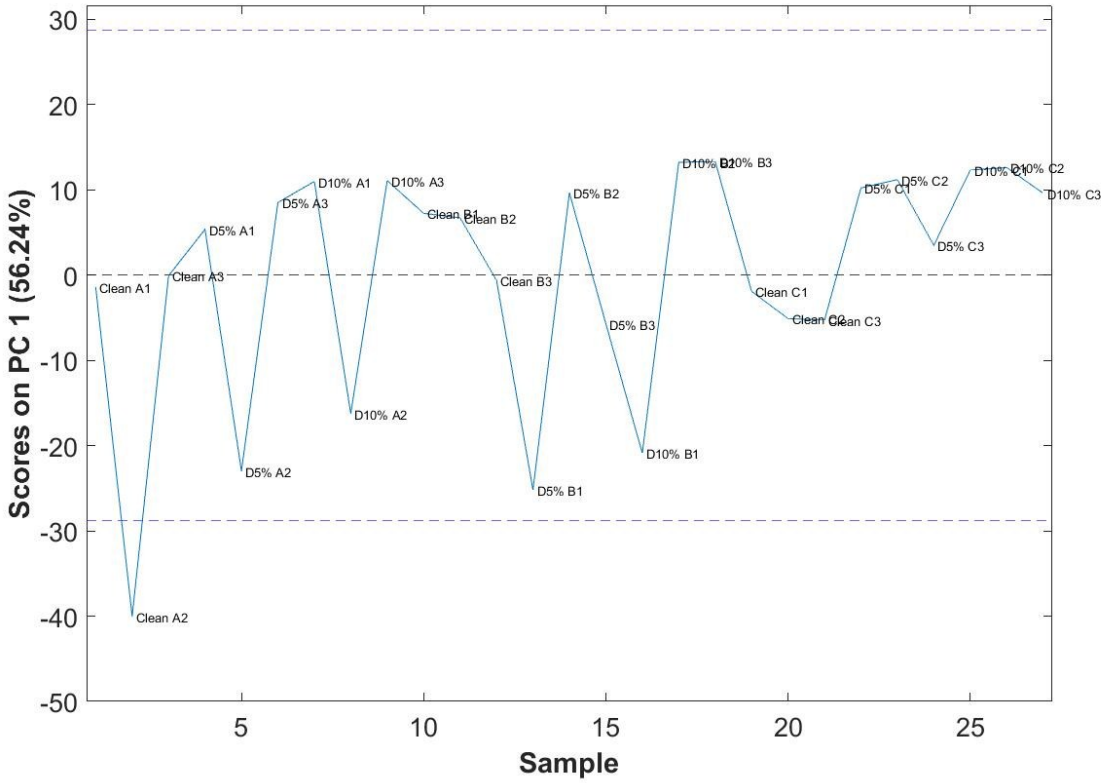
Supporting Information Figure 2. Dendrogram from a cluster analysis of the 0 %, 5 %, and 10 % electropherograms under consideration in this study. 'Clean', 'D5%', and 'D10%' represent the 0 %, 5 %, and 10% samples, and 'A', 'B', and 'C' represent the three subjects. The numbers after 'A', 'B', or 'C' represent replica runs. The data was pre-processed using range-selection followed by autoscaling.



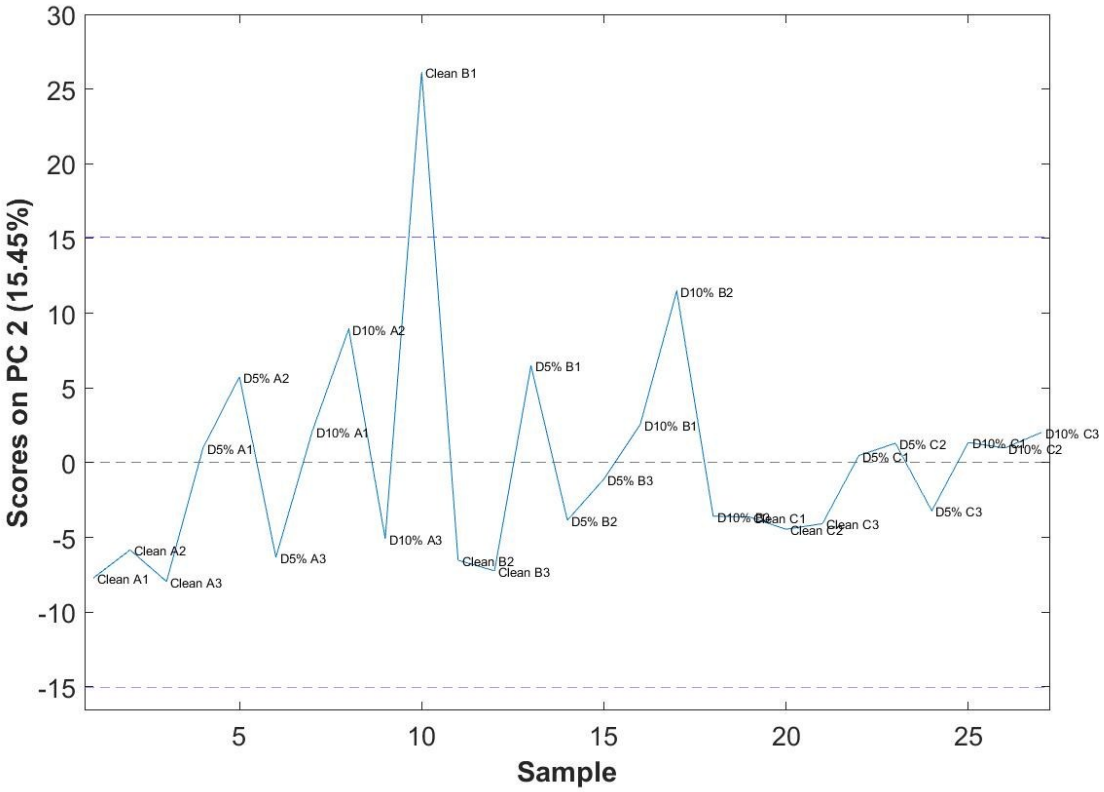
Supporting Information Figure 3. RMSEC (orange) and RMSECV (blue) plots from PCA calculations of the electropherograms from Subjects A, B and C for doped (5 % and 10 %) and undoped samples.



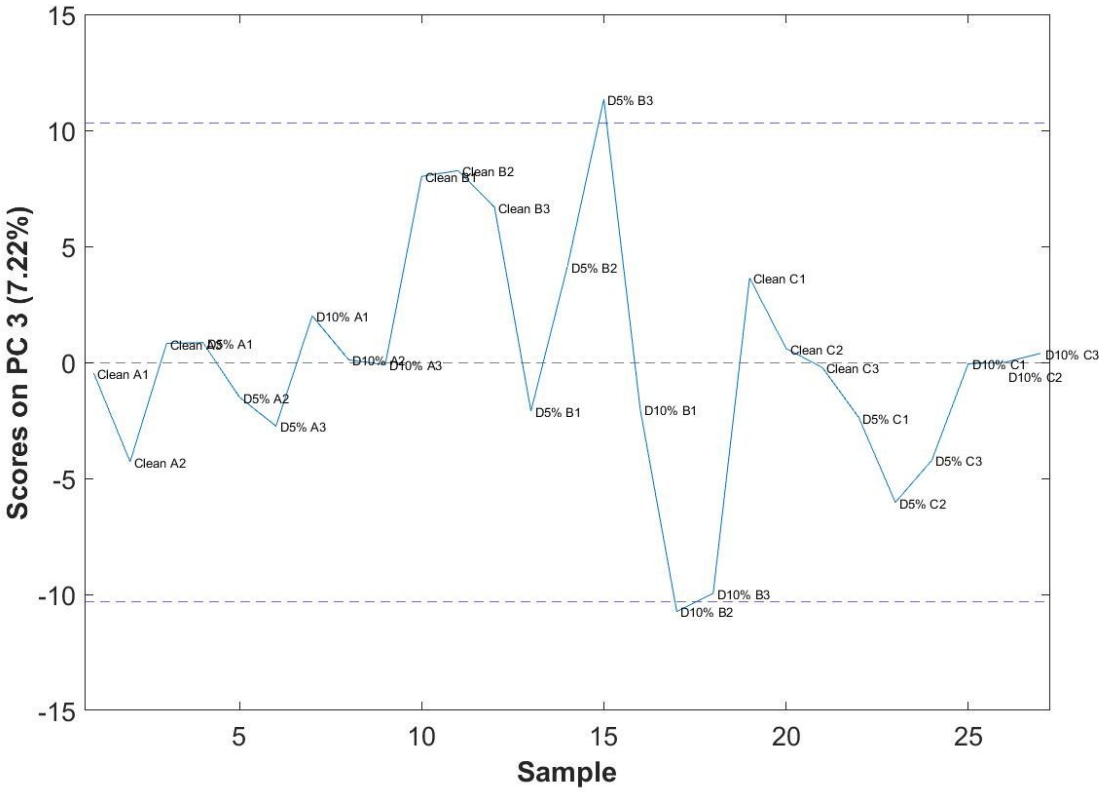
Supporting Information Figure 4. Scores on PC1 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 1-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



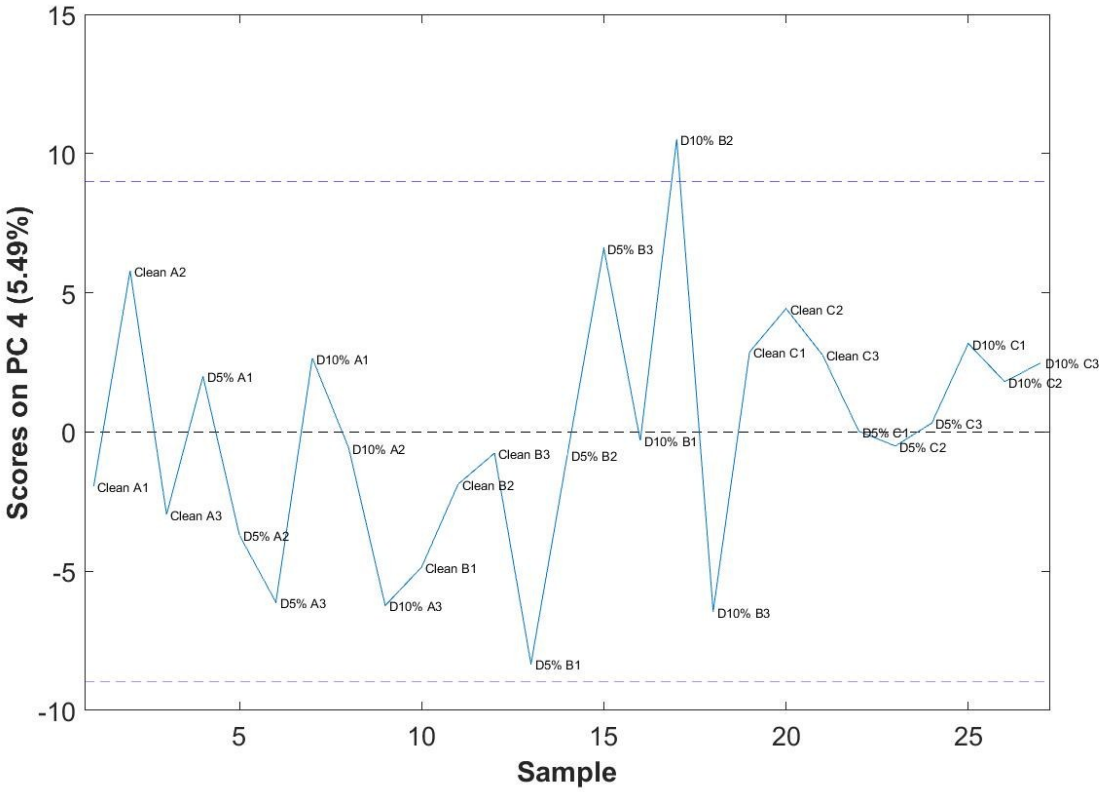
Supporting Information Figure 5. Scores on PC1 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



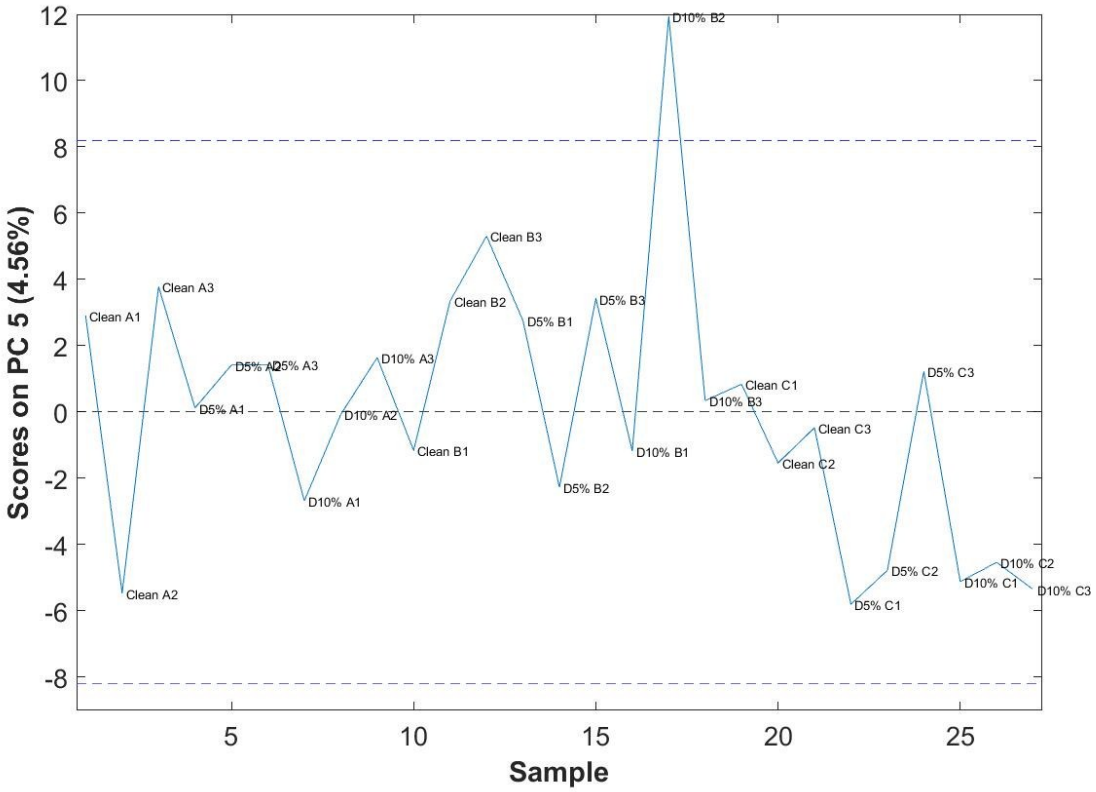
Supporting Information Figure 6. Scores on PC2 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



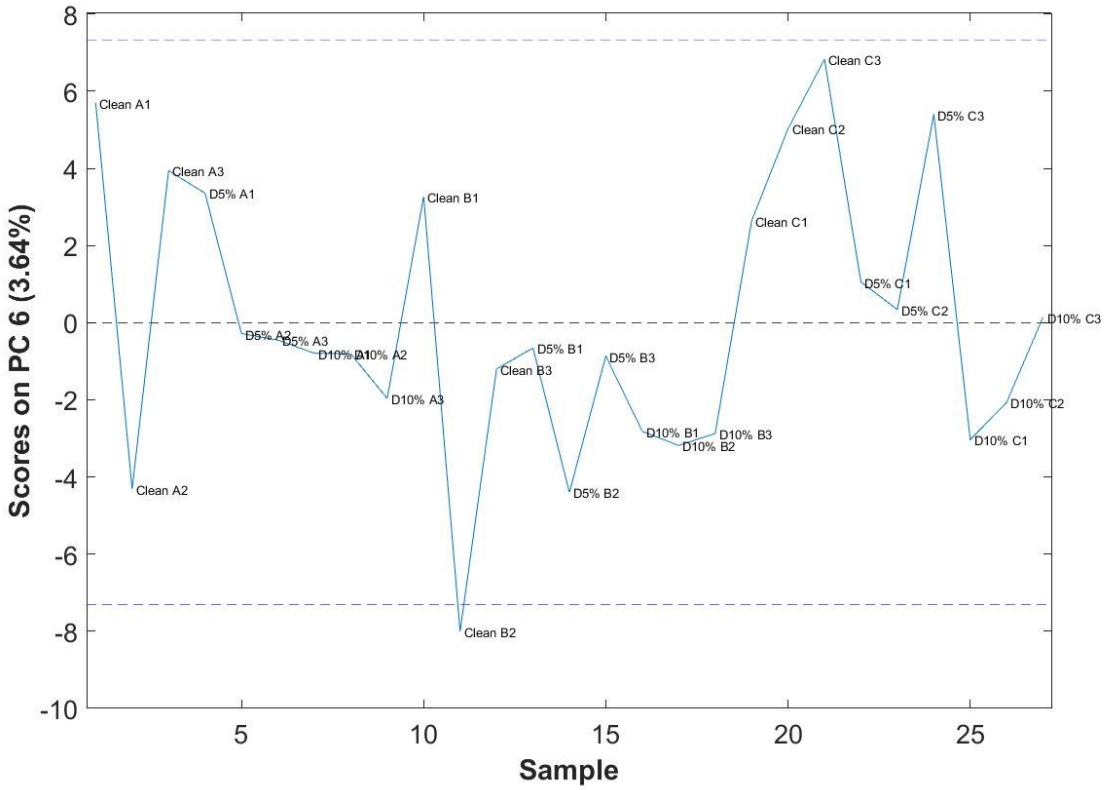
Supporting Information Figure 7. Scores on PC3 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



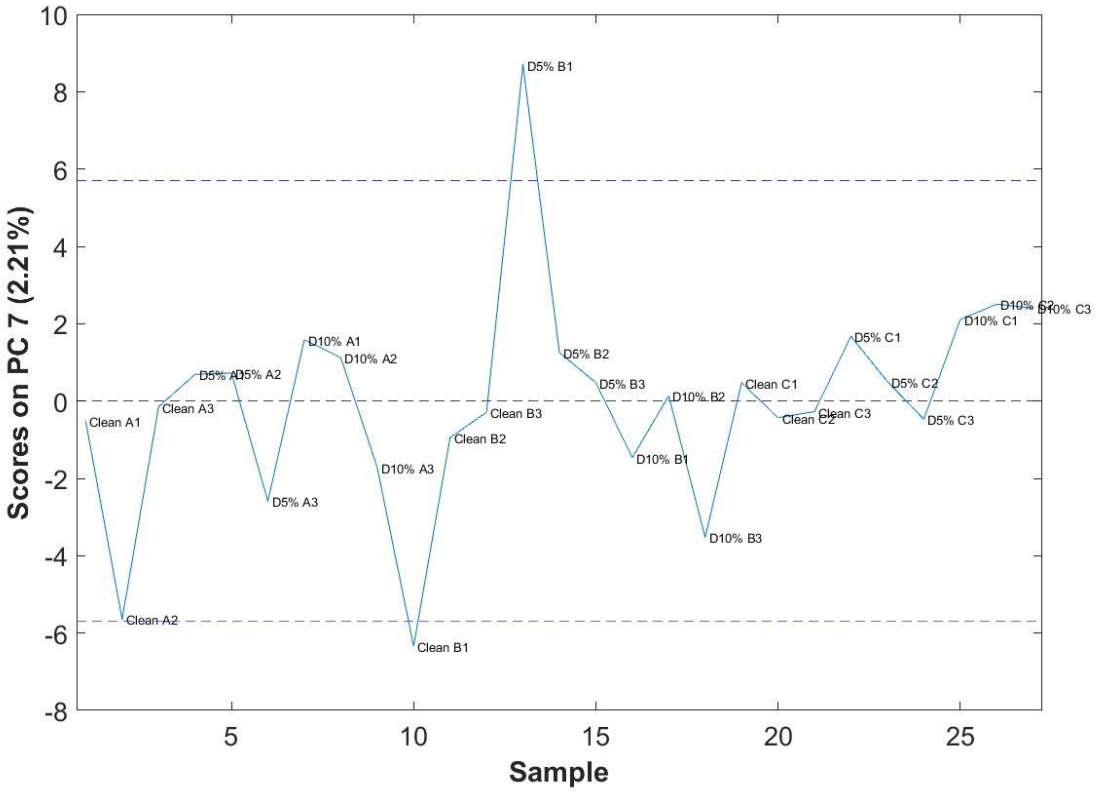
Supporting Information Figure 8. Scores on PC4 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



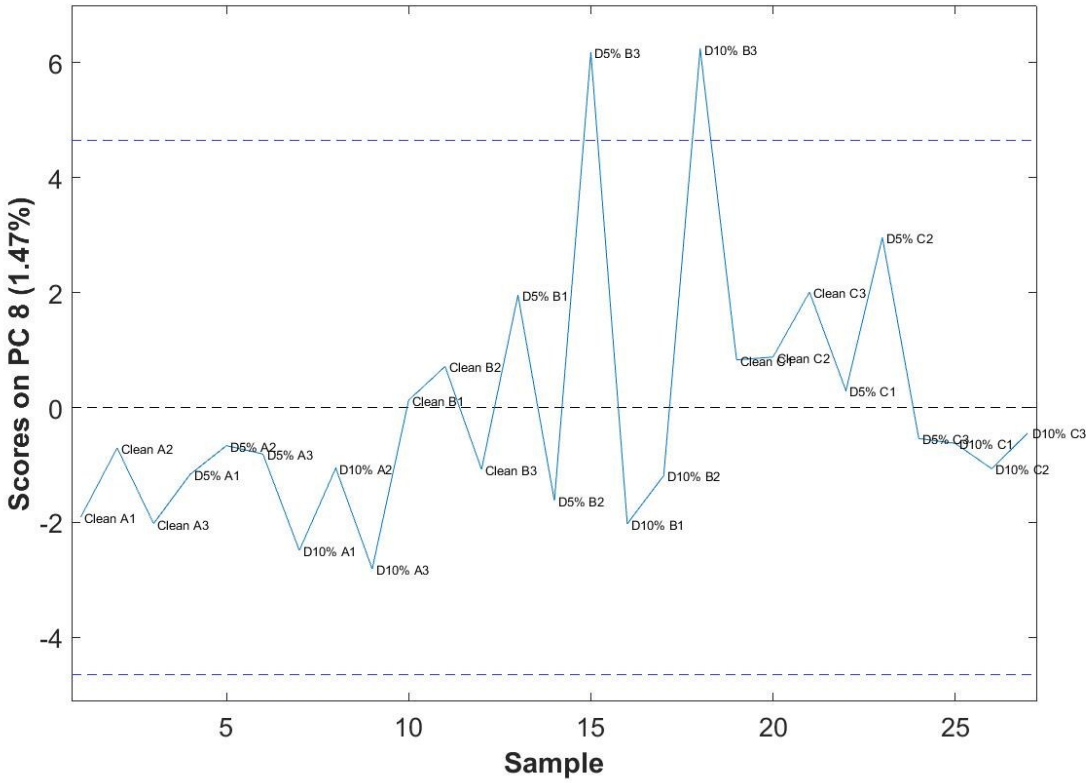
Supporting Information Figure 9. Scores on PC5 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



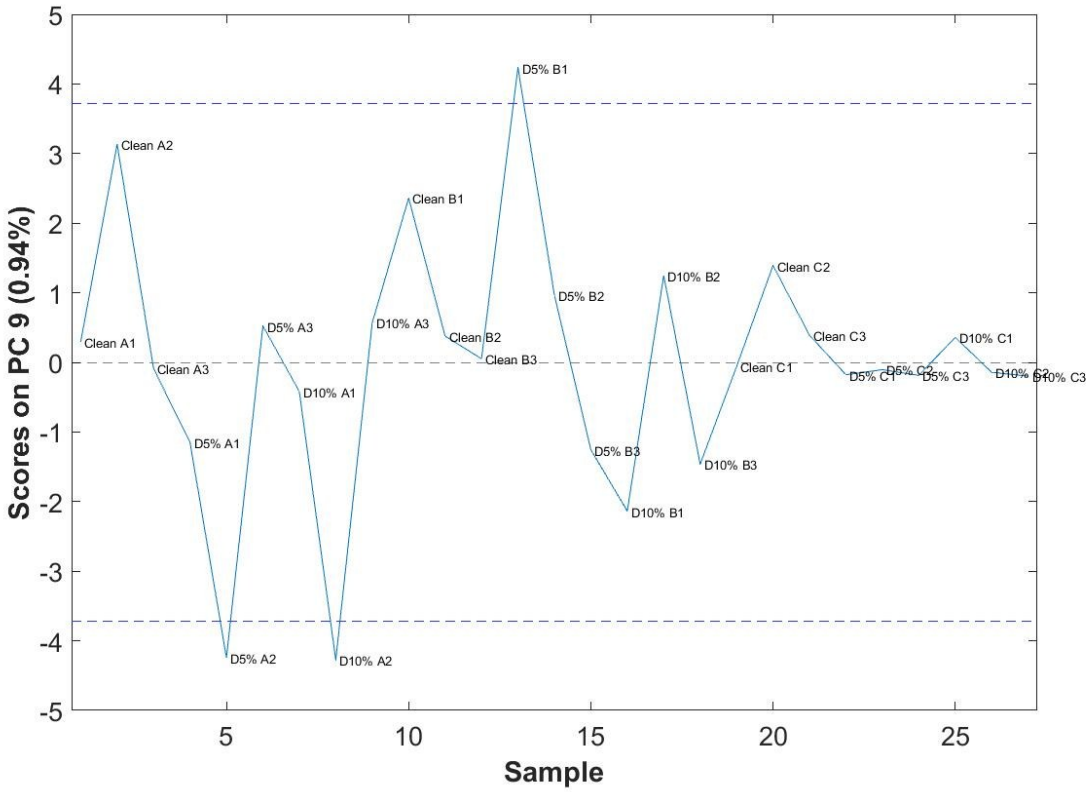
Supporting Information Figure 10. Scores on PC6 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



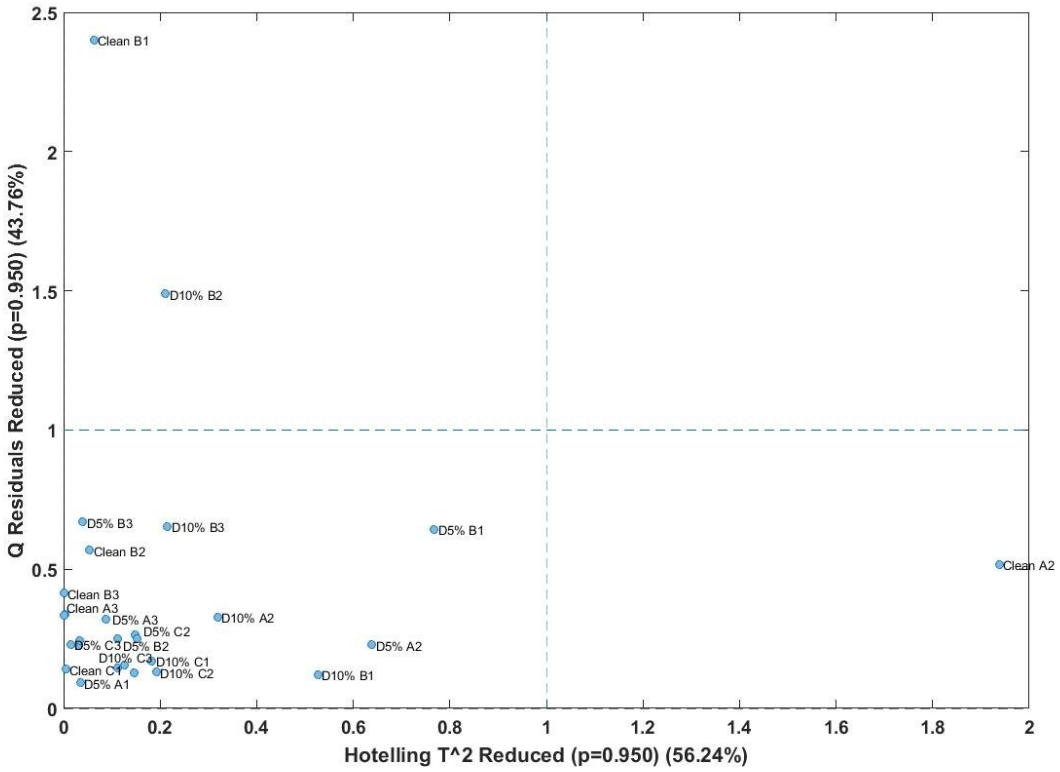
Supporting Information Figure 11. Scores on PC7 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



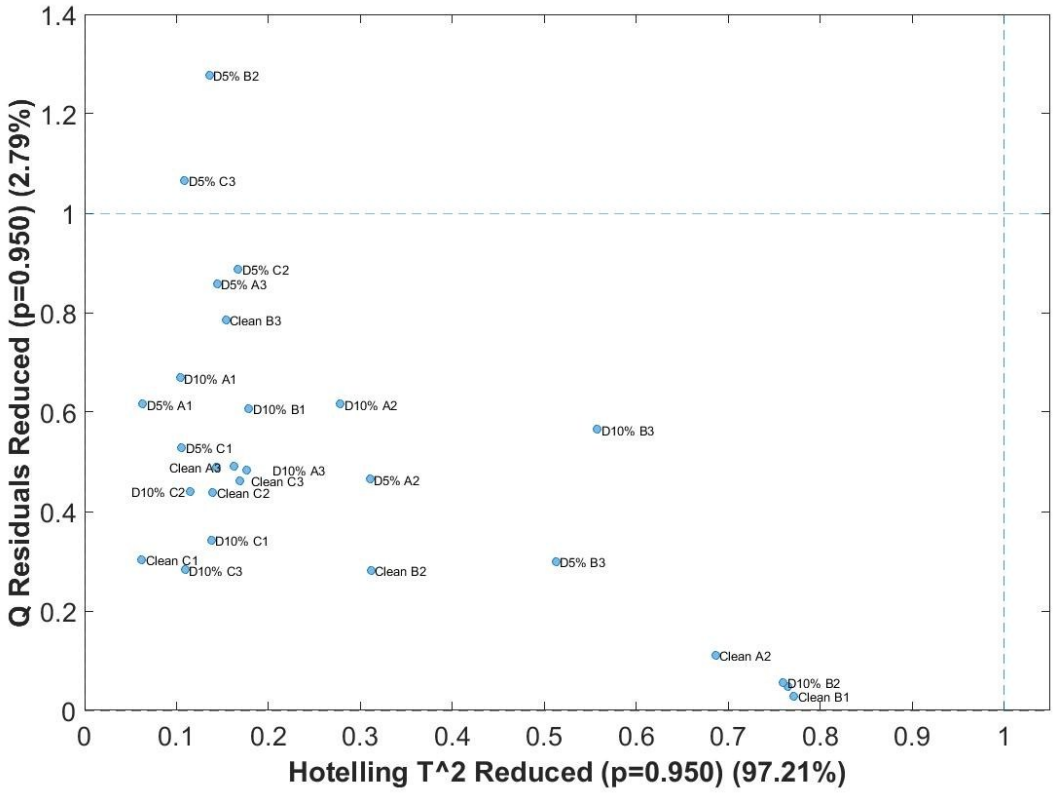
Supporting Information Figure 12. Scores on PC8 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



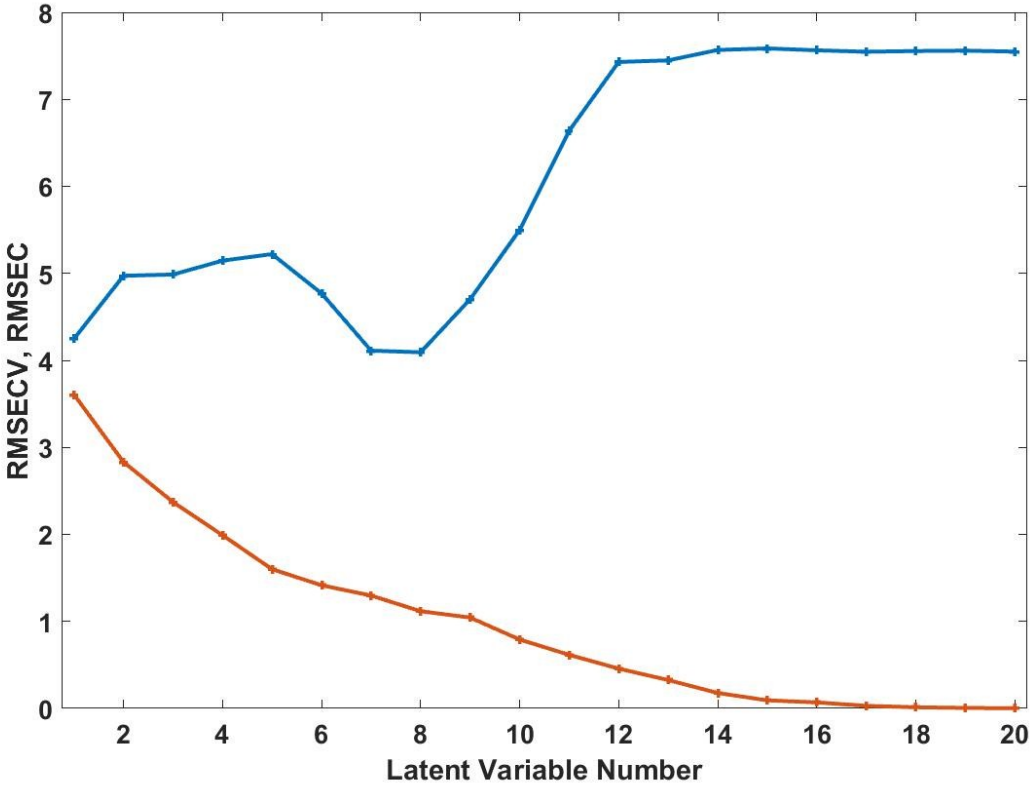
Supporting Information Figure 13. Scores on PC9 from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



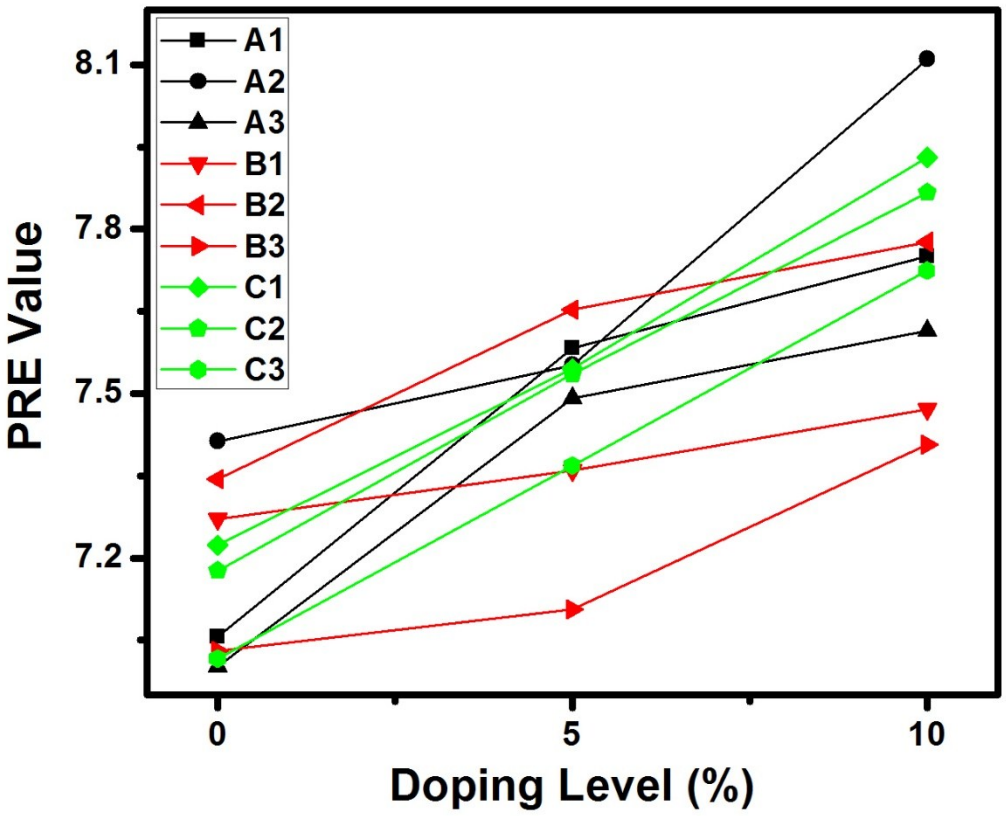
Supporting Information Figure 14. Hotelling T^2 vs Q residual plot from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 1-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



Supporting Information Figure 15. Hotelling T^2 vs Q residual plot from replicate runs of Subjects A, B and C for doped (5 % and 10 %) and undoped samples for a 9-Component PCA Model. Note: The undoped samples are labeled as 'Clean' and the replicate runs are labeled with the subject name and run number, i.e., A2 represents the second replicate run of Subject A.



Supporting Information Figure 16. RMSEC (orange) and RMSECV (blue) plots from PLS calculations of the electropherograms from Subjects A, B and C for doped (5 % and 10 %) and undoped samples.



Supporting Information Figure 17. Raw PRE values of electrophoretic separations of subjects A, B and C for various doping levels (0 %, 5 % and 10 %).