**Supplementary Material**

**Table S1:**  A summary explaining the abbreviations used in the statistical analysis

|  |  |
| --- | --- |
| **A** | number of components |
| **N** | number of samples |
| **R2X** | Sum of Squares of all the x-variables explained by the extracted components |
| **R2Y** | For PLS, Sum of Squares of all the y-variables explained by the extracted components |
| **Q2(Cum)** | Cumulative cross validated R2. |
| **eigenvalue** | An **eigenvalue** is a number, telling you how much variance there is in the data in that direction.  A larger **eigenvalue means** that that principal component explains a large amount of the variance in the data. |
| **PCA** | Principal component analysis; unsupervised MVA method suitable for data overview and identification of outliers. |
| **OPLS** | Orthogonal projections to latent structures; supervised MVA method suitable for variable selection or classification of 2 groups. |
| **Model** | The plane (or hyperplane) to which the data are projected, consisting of as many dimensions as principal components extracted. |
| **Principal component** | The coordinates of the original observations following reduction in dimensionality to a few latent variables. |
| **Loadings** | New values representing the variables in the model plane. Each original variable is represented as a point in the *loadings plot* |
| **Hotellings T2** | Multivariate generalization of the 95% confidence interval, can be utilized to identify outliers |
| **p(corr)** | Loadings scaled as a correlation coefficient (ranging from −1.0 to 1.0) between the model and original data. |
| ***R*2** | The fraction of the original data explained by the model (*R*2 = 1.0 explains 100% of the data). Measure of the overall fit of the model. |
| ***Q*2** | The fraction of the original data explained by the cross-validated model. Measure of the ability of the model to predict a new dataset. |
| **CV-ANOVA** | Cross-validated analysis of variance; provides a *p*-value indicating the level of significance of group separation in OPLS analyses. Based on a cross-validated model. |
| **RMSEP** | Measures the predictive power of the model. |
| **RMSECV** | A predictivity measure for the model that summarizes the cross-validation residuals of the observations in the workset. |
| **RMSEE** | A performance measure for the model that relates to the unit of the observed Y. |
| **VIP** | Variable importance in the projection; ranking of the original variables according to their individual contribution to the model. |

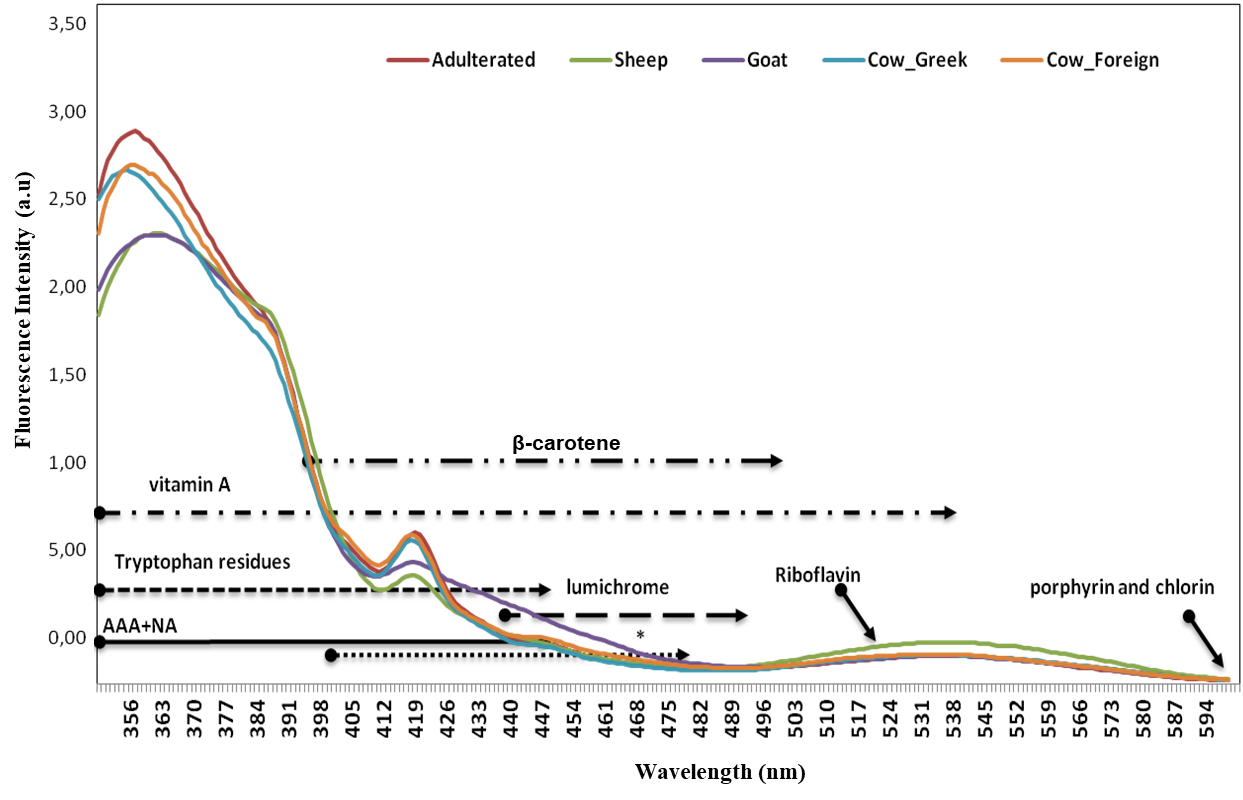
**Table S2:** A summaryof the PCA models involved in the optimal selection of Δλ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Δλ** | **R2X(cum)** | **Q2(cum)** | **R2X(cum) -Q2(cum)** | **Eigenvalue** |
| **100** | **0,84** | **0,83** | **0,01** | **PC1: 58,9, PC2:20,5** |
| 90 | 0,76 | 0,63 | 0,13 | PC1: 27,2, PC2:12,2 |
| 80 | 0,8 | 0,75 | 0,05 | PC1:38,9, PC2: 16,4 |
| 70 | 0,79 | 0,74 | 0,05 | PC1: 19,1, PC2:9,46 |
| 60 | 0,75 | 0,6 | 0,15 | PC1: 26,8, PC2: 13,3 |
| 50 | 0,8 | 0,78 | 0,02 | PC1: 50,6, PC2: 19,4 |
| 40 | 0,8 | 0,78 | 0,02 | PC1: 50,6, PC2: 19,4 |
| 30 | 0,79 | 0,71 | 0,08 | PC1: 55,6, PC2:23,8 |
| 20 | 0,81 | 0,78 | 0,03 | PC1: 57,2, PC2: 13,7 |
| 10 | 0,83 | 0,8 | 0,03 | PC1: 28,2, PC2: 9,95 |

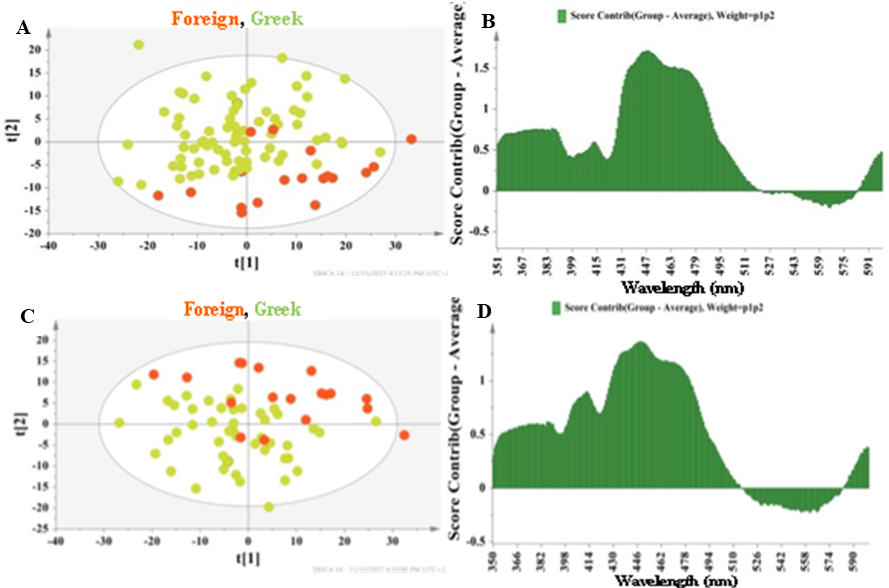
**Table S3** A summaryof the performance of the supervised models

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **RMSEE** | **RMSEcv** | **RMSEP** | **R2** | **Q2** | **R2-Q2** | *p*-value | **AUROC** |
| **Figure 3.A** | 0,263542 | 0,263542 | 0,600218 | 0,87 | 0,72 | 0,15 | 2,91E-08 | foreign 0,89, greek 0,73 |
| **Figure 3.B** | 0,246911 | 0,250762 | 0,604411 | 0,83 | 0,7 | 0,13 | 2,55E-11 | foreign 0,89, greek 0,74 |
| **Figure 3.C** | 0,29171 | 0,29284 | 0,98013 | 0,79 | 0,54 | 0,25 | 7,98E-11 | foreign 0,85, greek 0,61 |
| **Fig 5A** | 0,22584 | 0,22805 | 0,518635 | 0,98 | 0,76 | 0,22 |  | sheep 0,99; cow 0,96; goat 0,98 |
| **Fig7** | 0,26126 | 0,27743 | 0,2826 | 0,77 | 0,3 | 0,47 | 0,007143 | fraud 0,9; cow 0,75 |

**Figure S1:** Indicative SyFS spectrum for Δλ=100 nm of cow, sheep, goat milk spectra and adulterated and foreign milk samples’ with annotations on compounds within the fluorescent areas.



**Figure S2Α**. PCA model N=94, A=2, R2X(cum)= 0.77, Q2(cum)= 0.58, (Greek: Green circles) (Foreign: Orange circles).**B**. Contribution plot of Greek vs Foreign cow milk samples depicting the variables responsible for the differentiation between all the Greek cow milk samples and all foreign cow samples



## Fig.S3: Permutation testing for the OPLS-DA model in Figure 3.A

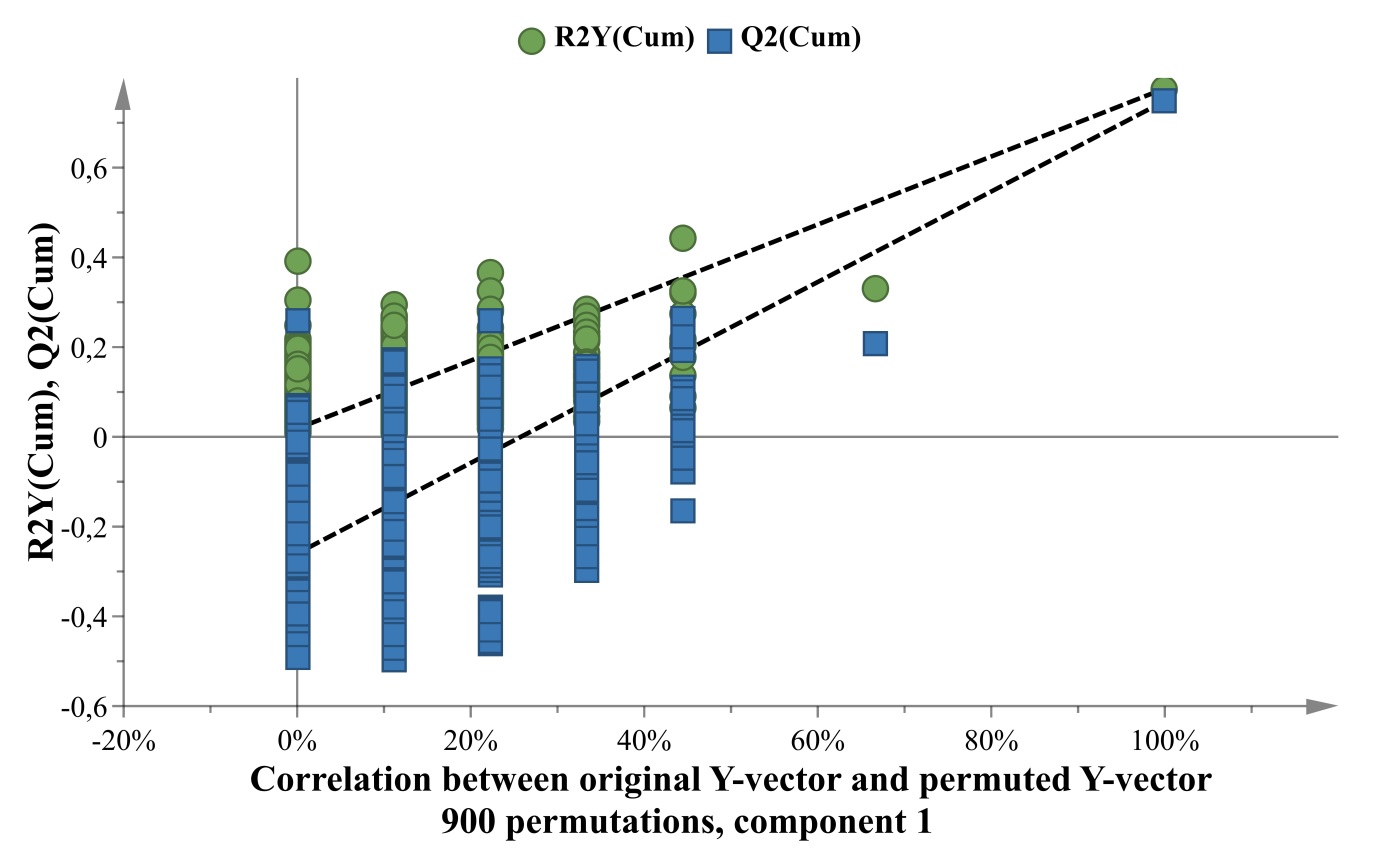
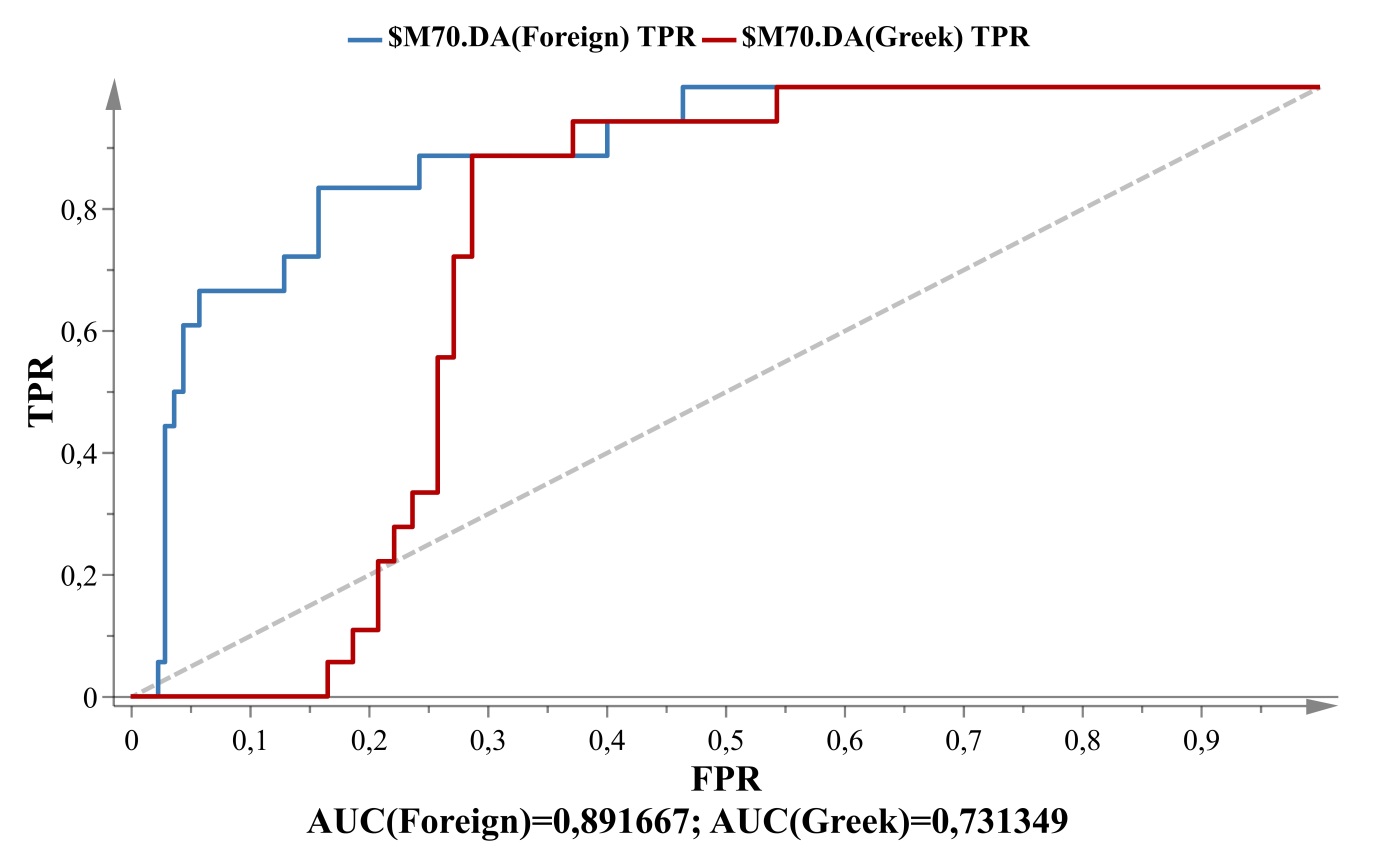
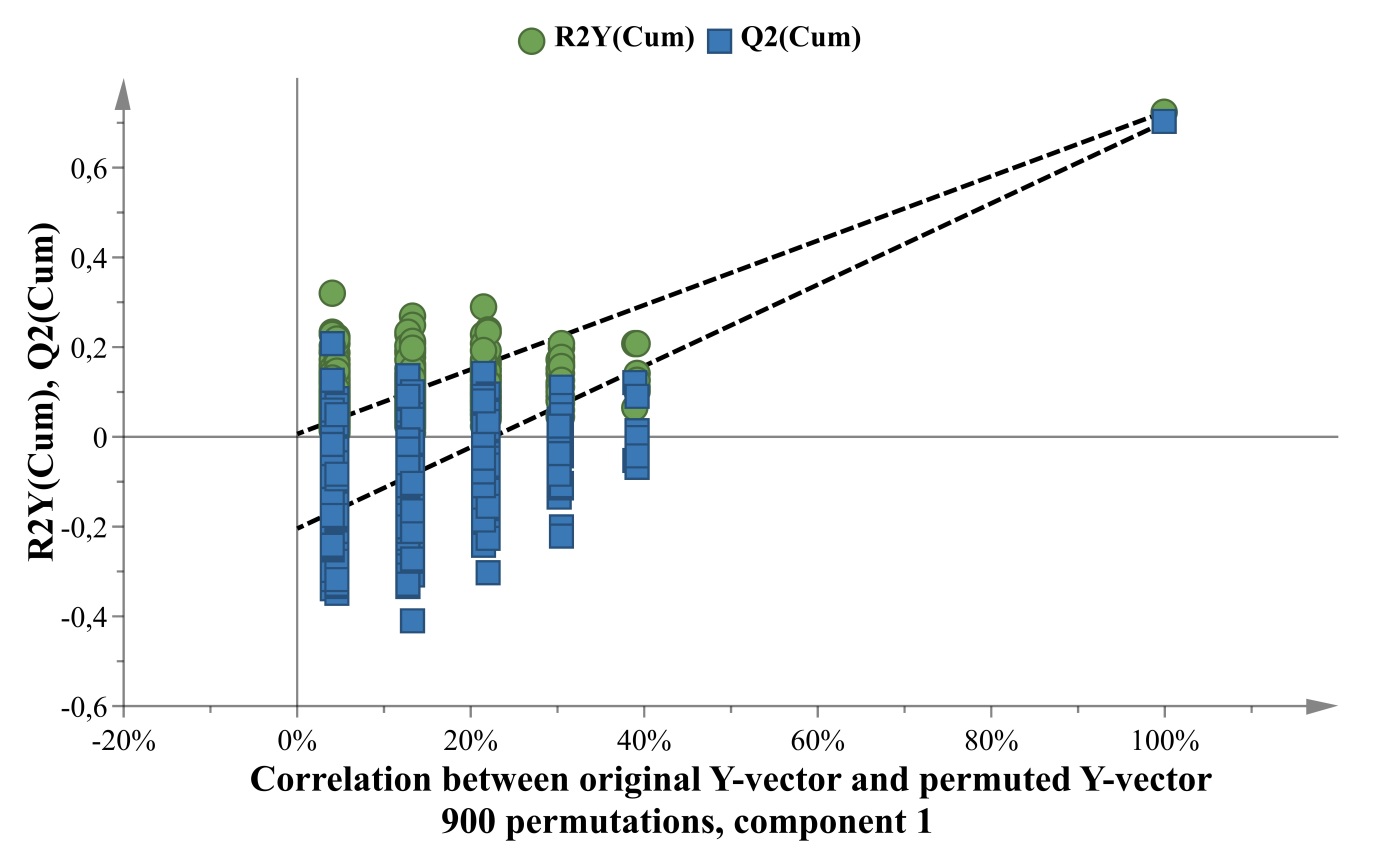
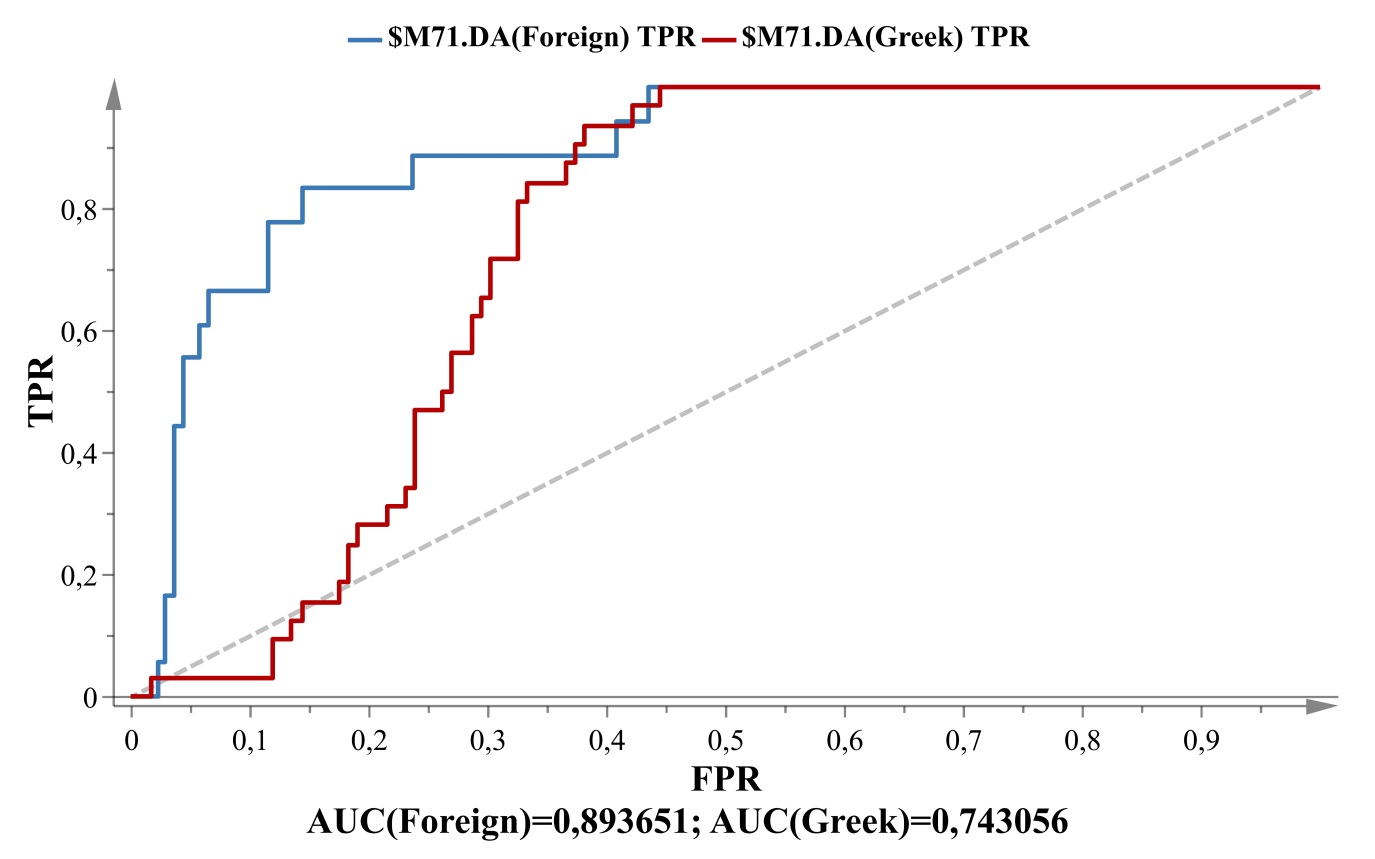
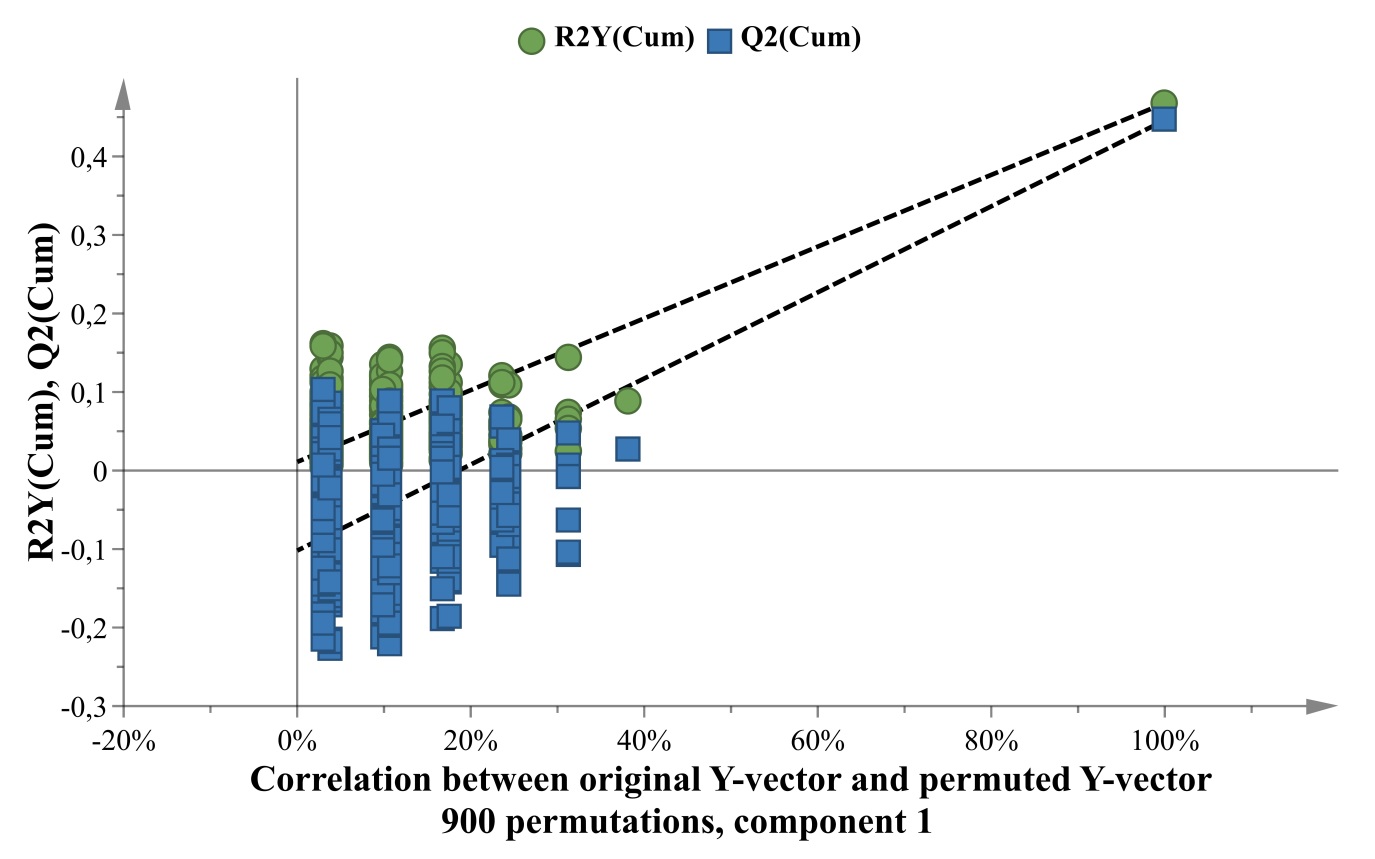
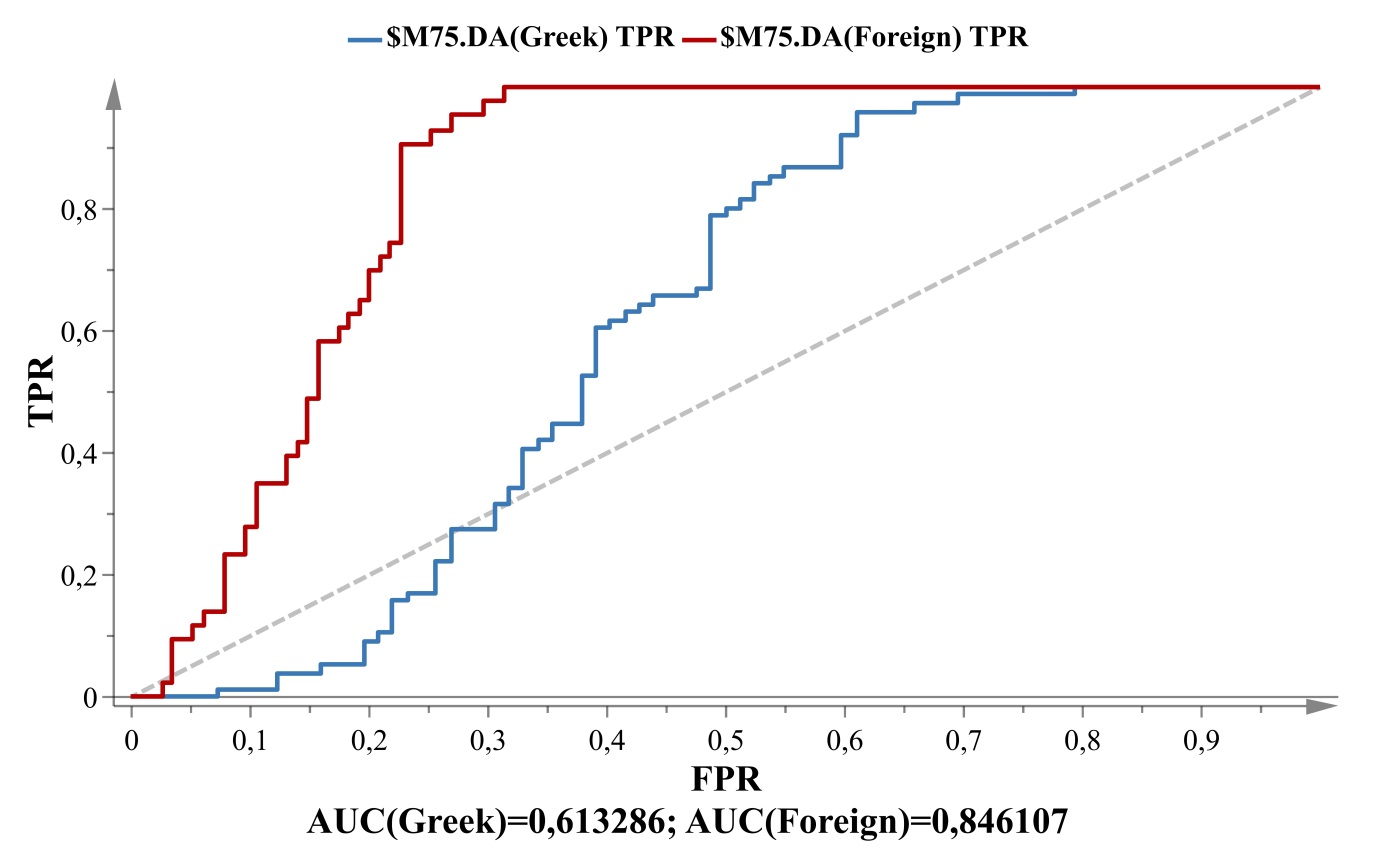


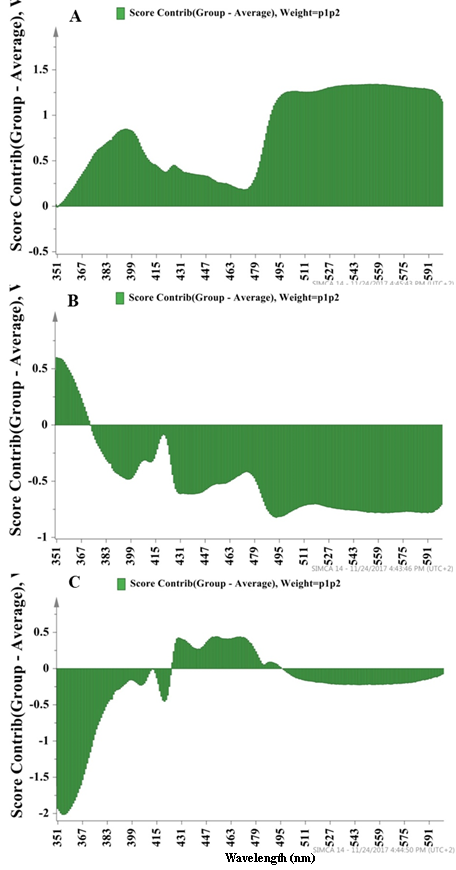
Fig.S4: **Permutation testing for the OPLS-DA model in Figure 3.B**



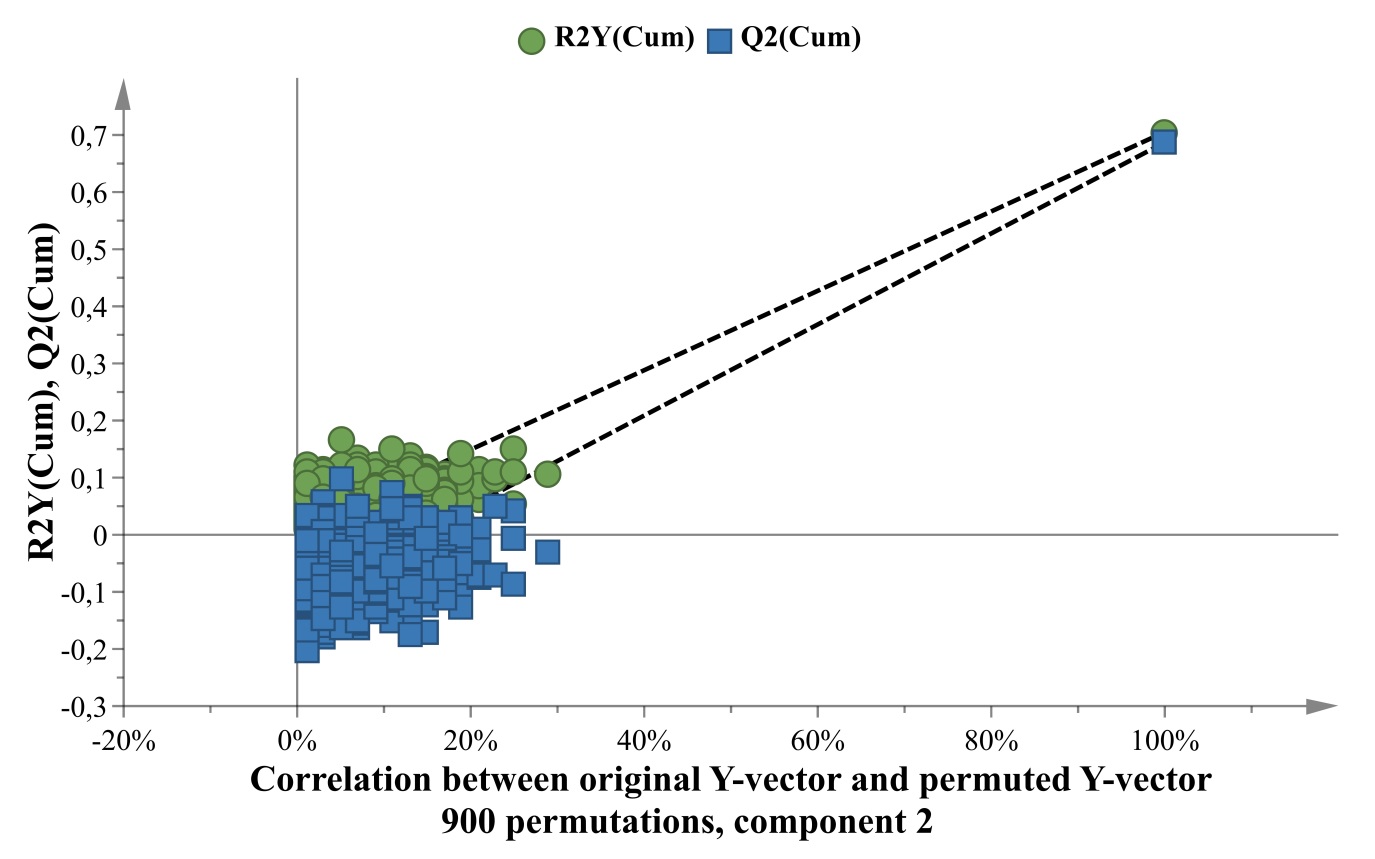
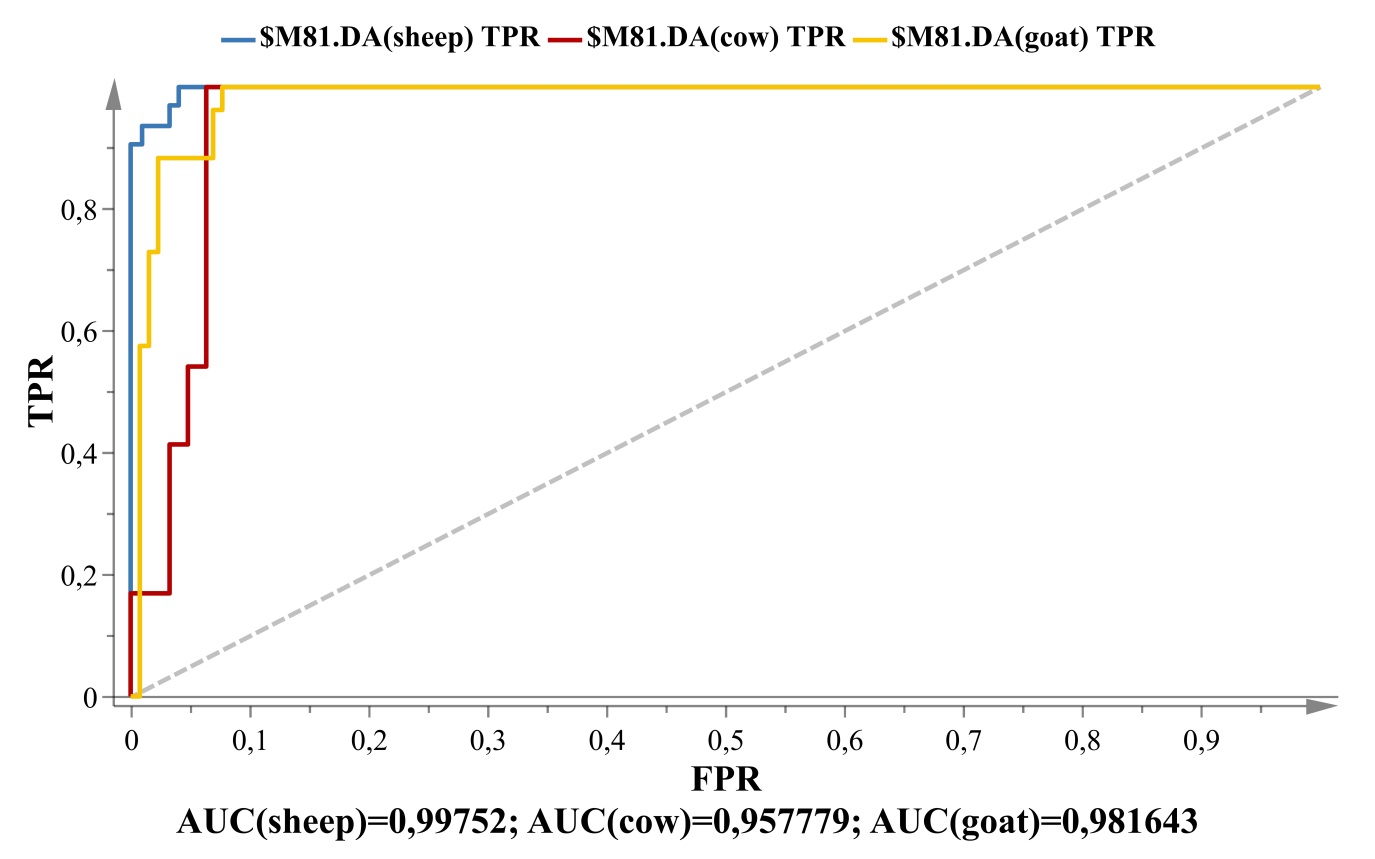
## Fig.S5: Permutation testing for the OPLS-DA model in Figure 3.C



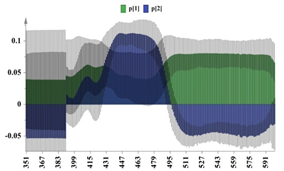
**Figure S6:** Contribution plot of sheep samples regarding PCA model of Figure 4, **B**. Contribution plot of cow samples regarding PCA model of Figure 5 , **C**. Contribution plot of goat samples regarding PCA model of Figure 5.



## Fig.S.7: Permutation testing for the OPLS-DA model in Figure 5



**Figure S8** Loading plot of the first two components regarding the PCA model of Figure 6



**Figure S9** : **Permutation testing for the OPLS-DA model in** Figure 6

