Supporting Information Available

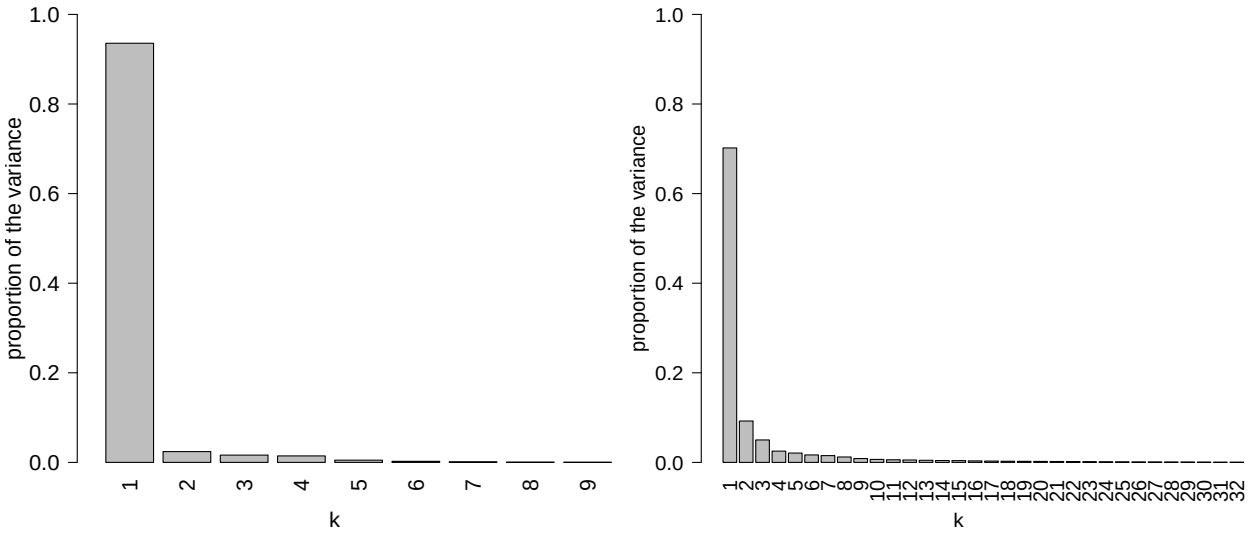


Figure I: Proportion of the variance explained by the principal components of the control metabolites in the (a) multi-site study ($k = 2, \ldots, 5$ were explored), and (b) multi-cohort study ($k = 6, \ldots, 12$ were explored)

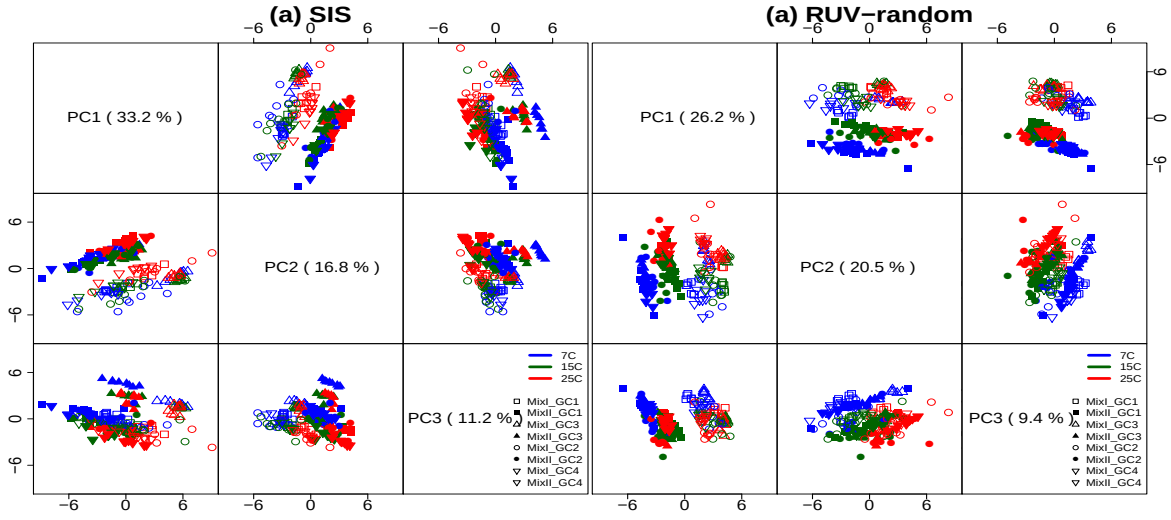
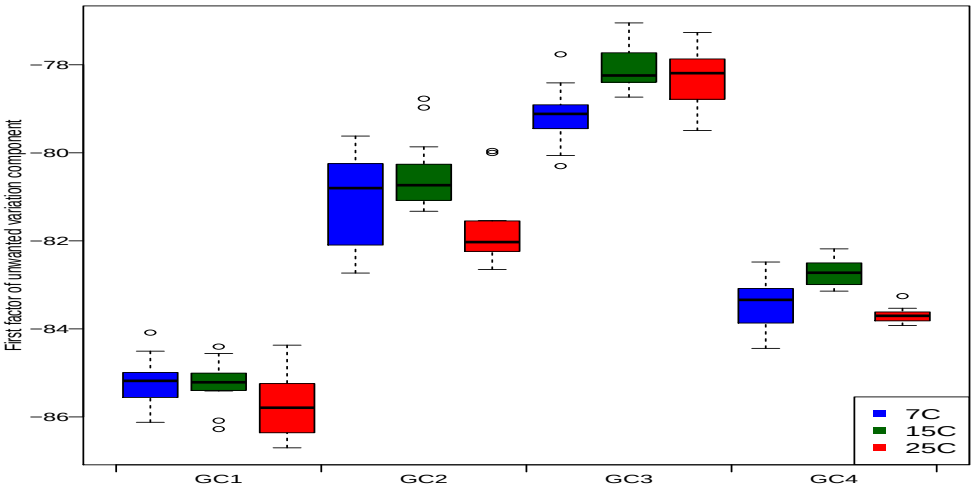
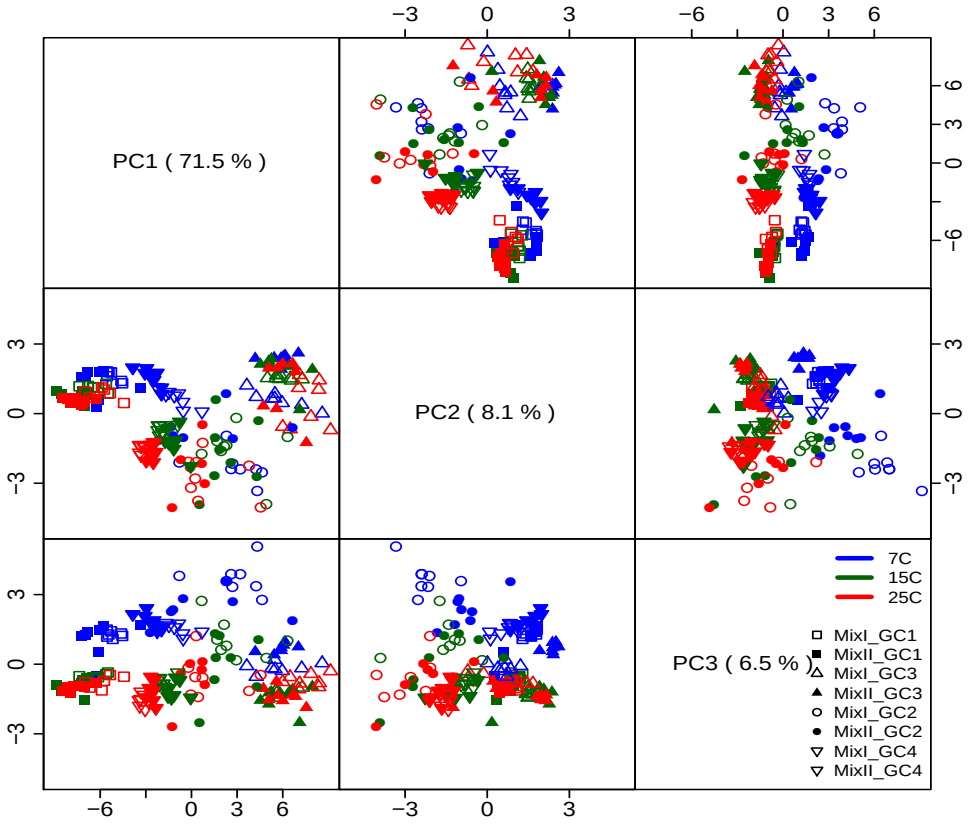


Figure II: Multi-site study: The first three principal components of the data normalized by (a) SIS (b) RUV-random without iteration ($k = 3$,$\lambda = 0.03$). The colours and shapes indicate different temperatures and instruments respectively, and Mix I and Mix II samples are shown by the hollow and solid points respectively.



(a) Boxplots of the samples showing the unwanted variation captured by the estimated first factor of unwanted variation



(b) The first three principal components of the unwanted variation removed by the RUV-random method improved for clustering

Figure III: Multi-site study: Figures showing the (a) boxplots of the samples contributing to the estimated first factor of unwanted variation, and the (b) first three principal components of the estimated unwanted variation component which was removed by the RUV-random method improved for clustering ($k = 8$, $\lambda = 1.43$). The colours and shapes indicate different temperatures and instruments respectively, and Mix II samples are shown by the solid points.

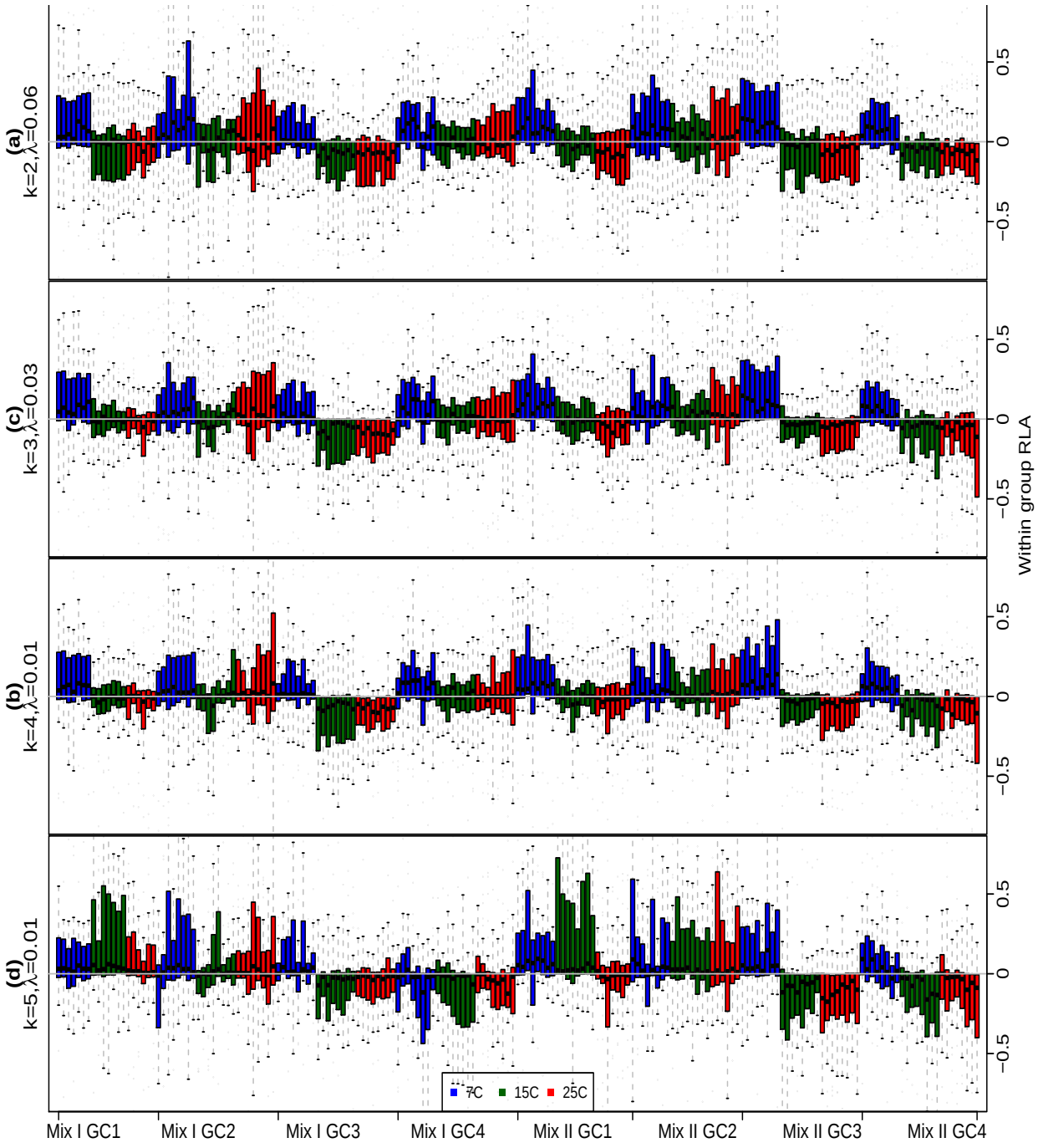


Figure IV: Multi-site study: Within-group RLA plots of the RUV-random normalized data as both λ and k were jointly varied in the RUV-random method.

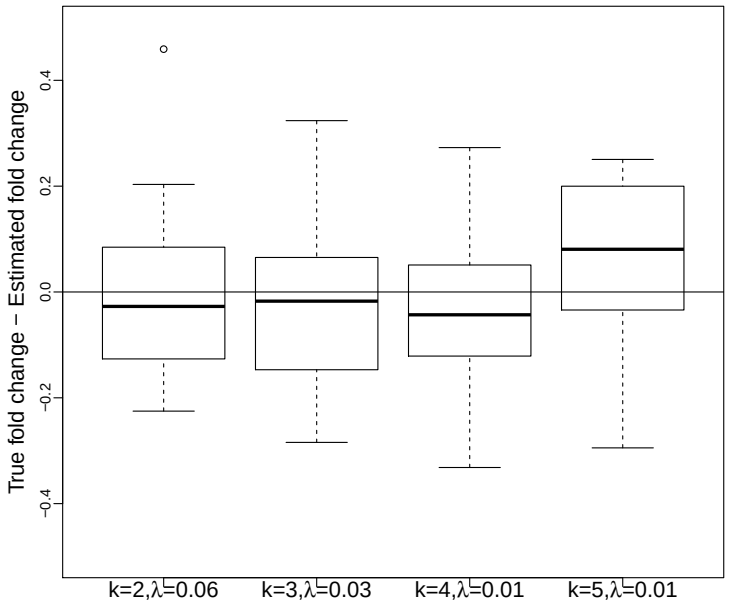


Figure V: Multi-site study: Boxplots of the differences between true and estimated fold changes, both λ and k were jointly varied in the RUV-random method.

k, λ	Number of misclassified Metabolites
2, 0.064	3
3, 0.034	4
4, 0.015	2
5, 0.008	2

Table I: Multi-site study: The number of misclassified metabolites obtained from the hierarchical cluster analysis, as both λ and k were jointly varied in the RUV-random method.

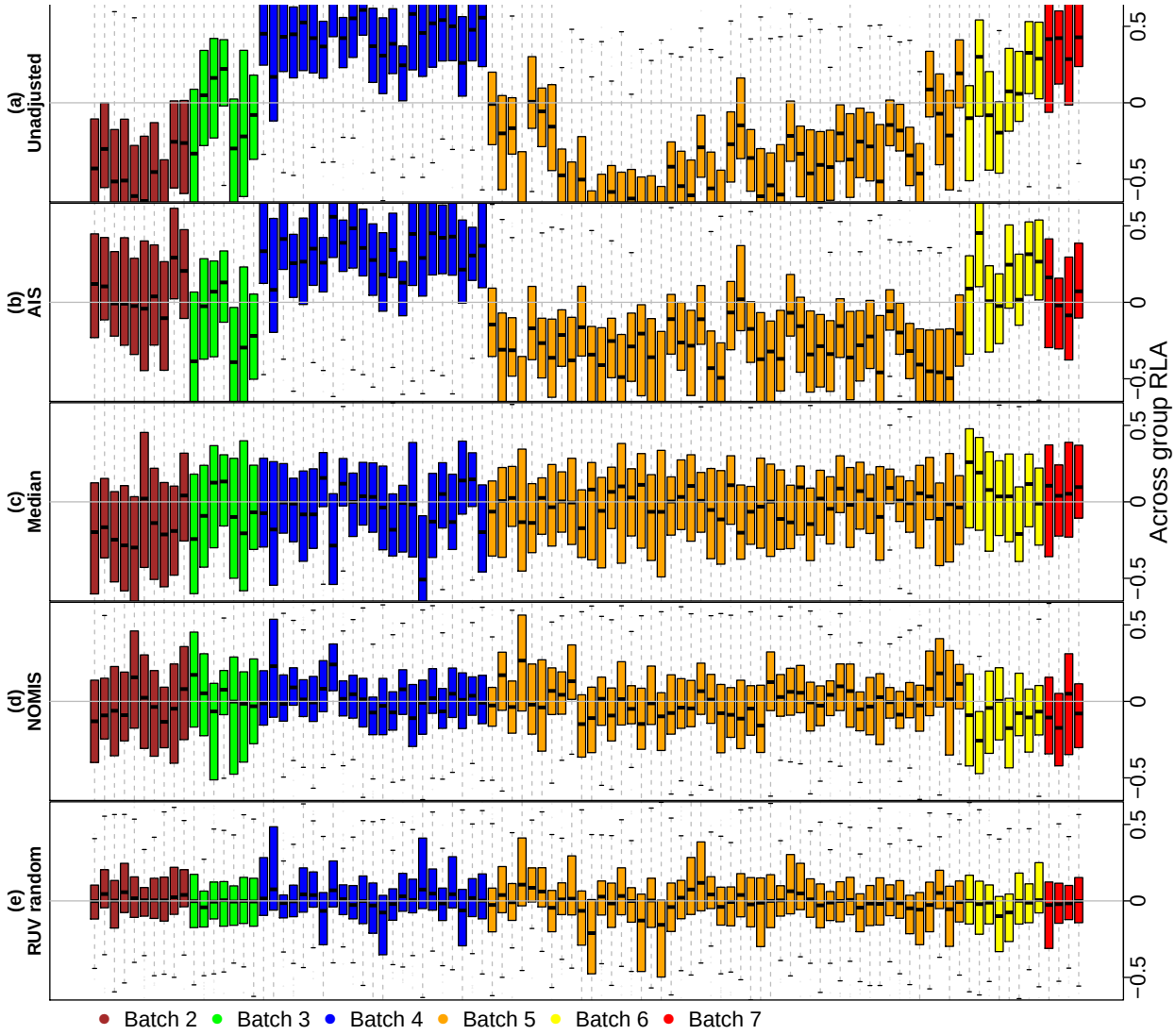


Figure VI: Multi-cohort study: Across-group RLA plots of the unadjusted and normalized data (For clearer visualisation, only one hundred randomly chosen samples are shown here, in run order). Colours indicate different batches.

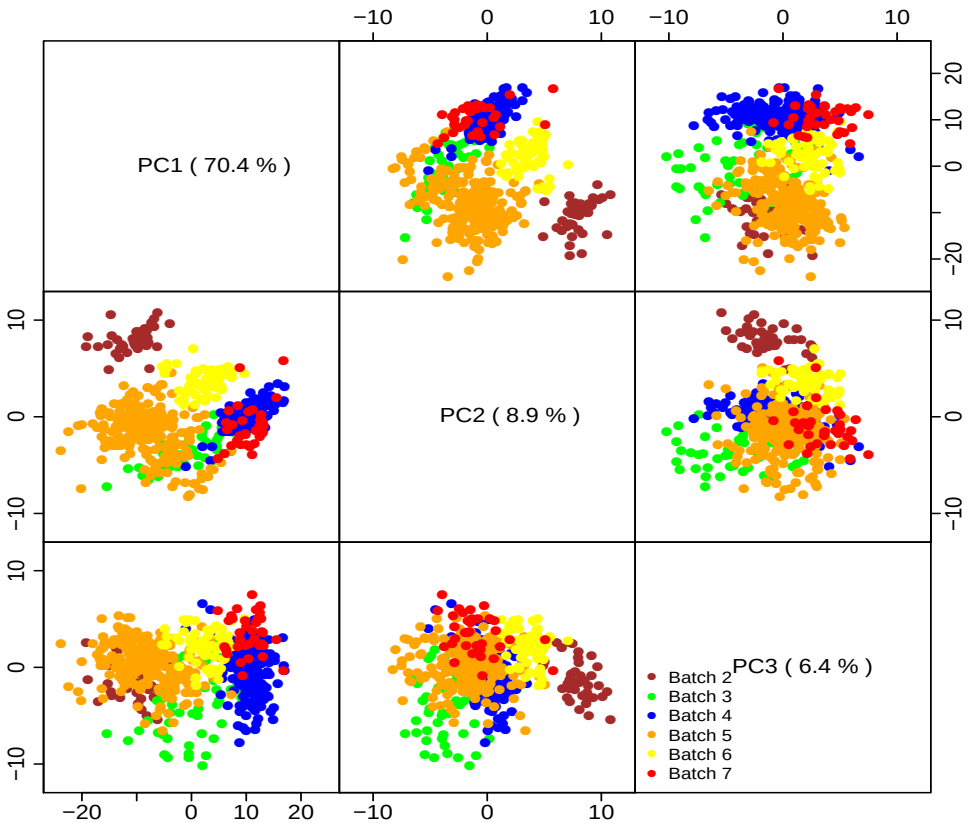


Figure VII: Multi-cohort study: The first three principal components of the unwanted variation removed by the RUV-random method ($k = 10$, $\lambda = 0.31$). Colours indicate different batches.

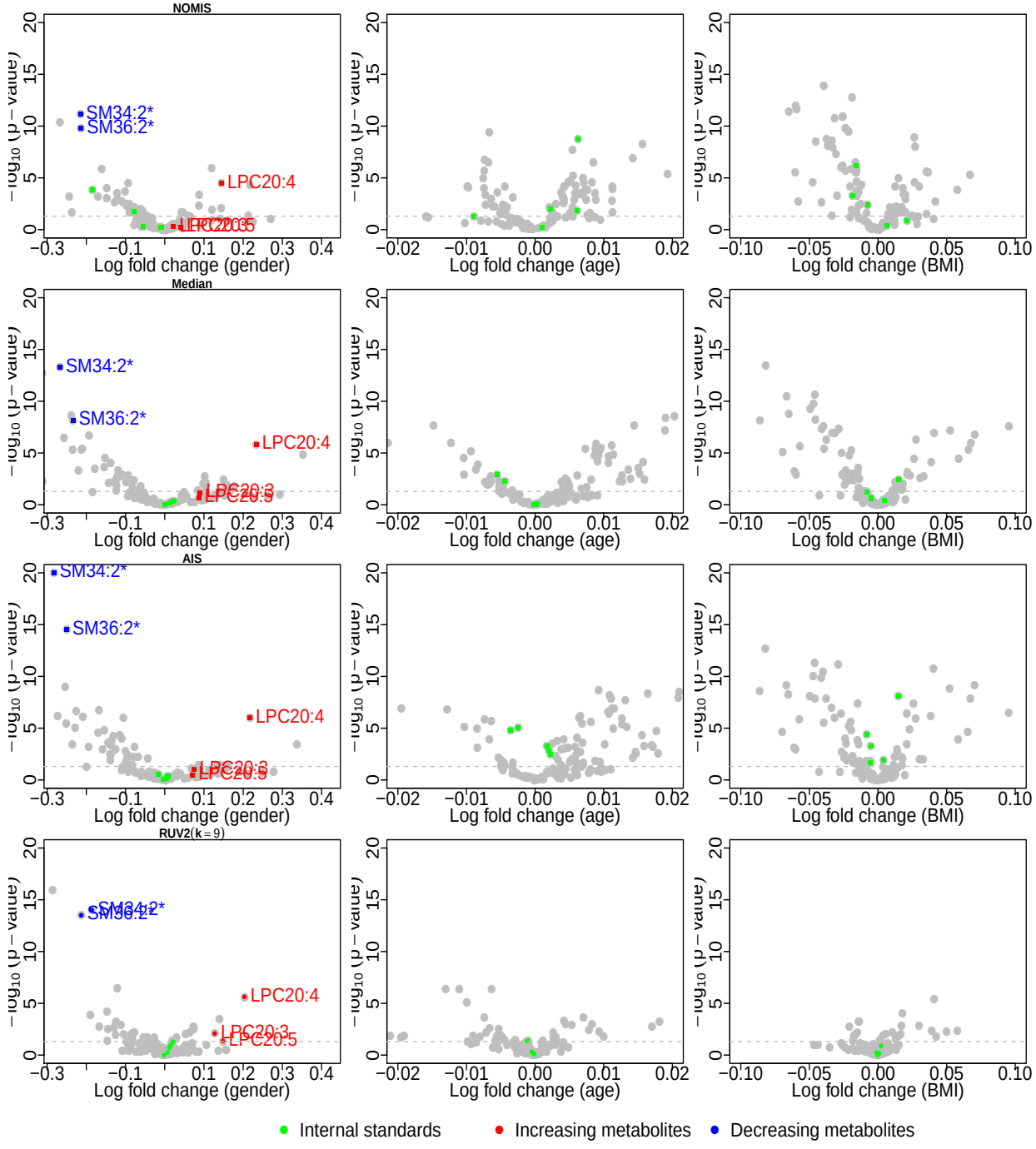


Figure VIII: Multi-cohort study: Volcano plots for the unadjusted and normalized data. The application-specific RUV-2 method[23,24] is presented for comparison.

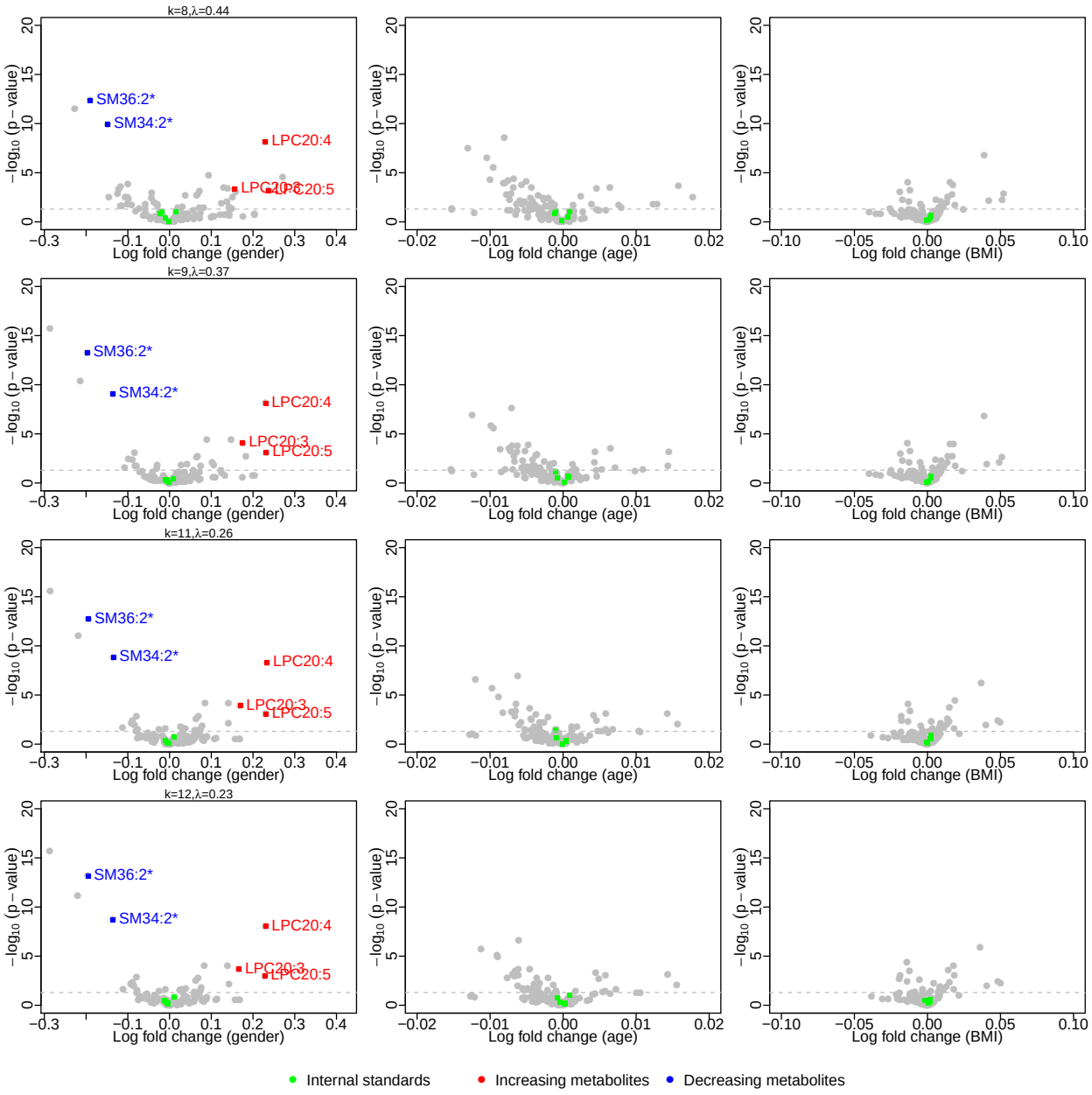


Figure IX: Multi-cohort study: Volcano plots obtained as both λ and k were jointly varied in the RUV-random method.

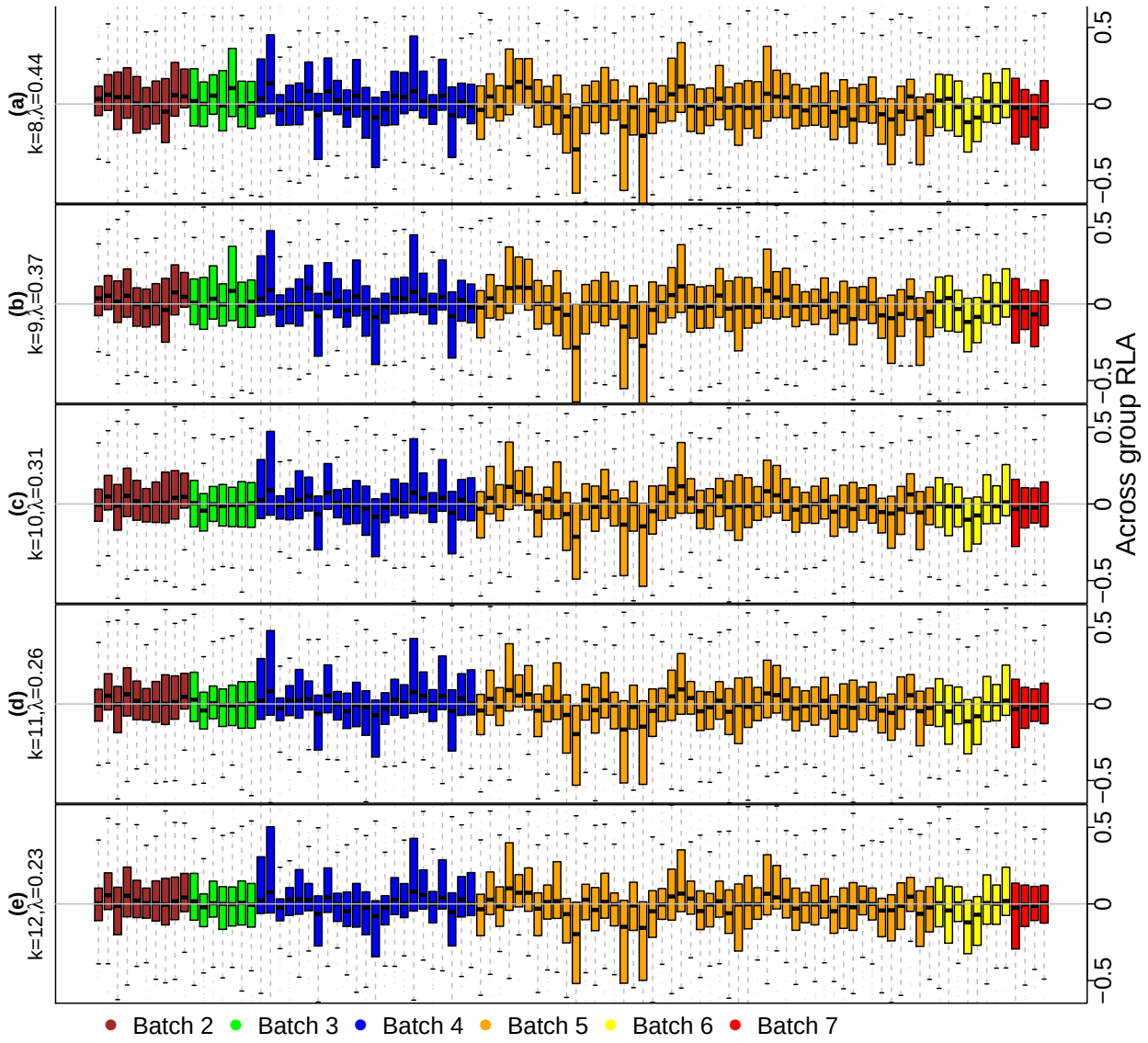


Figure X: Multi-cohort study: Across-group RLA plots as both λ and k were jointly varied in the RUV-random method. One hundred randomly chosen samples are shown here in run order, for better visualisation. Colours indicate different batches.