

Table S1. Validation set accuracy ($a_{ave} \pm a_{std}$ %) ($a_{ave} \pm a_{std}$: average accuracy rate \pm standard error over 20 repeats). The data sets was preprocessed by normalization before modeling..

Data set	Type	PLS-DA	BSS	GA-PLS-DA	s-PLS-DA	EBSS
<i>Olive oil</i>	FTIR	91.3 \pm 3.2	92.6 \pm 2.8	91.1 \pm 3.0	95.1 \pm 3.1	96.6 \pm 2.9
<i>Red wine</i>	FTIR	57.1 \pm 8.9	58.9 \pm 11.1	57.9 \pm 11.5	66.8 \pm 9.8	68.6 \pm 11.4
<i>NIR Tablet</i>	NIR	84.8 \pm 2.2	84.0 \pm 3.8	86.7 \pm 4.0	88.3 \pm 2.9	87.5 \pm 4.0
<i>Raman</i>	Raman	83.0 \pm 4.7	75.4 \pm 6.8	75.3 \pm 6.9	78.8 \pm 4.9	85.7 \pm 5.8
<i>Tablet</i>						

Table S2. The number of selected variables ($n_{ave} \pm n_{std}$) ($n_{ave} \pm n_{std}$: average number of selected variable \pm standard error over 20 repeats). The data sets was preprocessed by normalization before modeling.

Data set	Type	PLS-DA	BSS	GA-PLS-DA	s-PLS-DA	EBSS
<i>Olive oil</i>	FTIR	570	10 \pm 4	33 \pm 8	65 \pm 21	9
<i>Red wine</i>	FTIR	842	13 \pm 12	32 \pm 15	52 \pm 30	23
<i>NIR tablet</i>	NIR	404	28 \pm 30	43 \pm 6	59 \pm 18	27
<i>Raman tablet</i>	Raman	3041	36 \pm 20	55 \pm 13	76 \pm 19	30

Table S3.

Validation set accuracy ($a_{ave} \pm a_{std}$ %) ($a_{ave} \pm a_{std}$: average accuracy rate \pm standard error over 20 repeats). The data sets was preprocessed by standard normal variate transformation (SNV) before modeling.

Data set	Type	PLS-DA	BSS	GA-PLS-DA	s-PLS-DA	EBSS
<i>Olive oil</i>	FTIR	94.0 ± 2.0	90.8 ± 3.3	92.1 ± 4.4	94.2 ± 2.2	94.7 ± 3.0
<i>Red wine</i>	FTIR	58.2 ± 10.4	58.2 ± 13.9	58.2 ± 11.0	67.5 ± 10.5	60.4 ± 10.5
<i>NIR Tablet</i>	NIR	84.7 ± 3.3	84.2 ± 3.3	82.2 ± 2.8	86.2 ± 3.1	89.0 ± 2.6
<i>Raman</i>	Raman	79.9 ± 4.2	74.2 ± 7.2	74.5 ± 7.3	72.6 ± 6.1	87.2 ± 6.2
<i>Tablet</i>						

Table S4. The number of selected variables ($n_{ave} \pm n_{std}$) ($n_{ave} \pm n_{std}$: average number of selected variable ± standard error over 20 repeats). The data sets was preprocessed by SNV before modeling.

Data set	Type	PLS-DA	BSS	GA-PLS-DA	s-PLS-DA	EBSS
<i>Olive oil</i>	FTIR	570	9 ± 3	37 ± 11	65 ± 23	11
<i>Red wine</i>	FTIR	842	11 ± 5	36 ± 12	57 ± 22	11
<i>NIR tablet</i>	NIR	404	15 ± 5	47 ± 4	64 ± 15	27
<i>Raman tablet</i>	Raman	3041	34 ± 16	62 ± 8	64 ± 19	29

Table S5.

Validation set accuracy ($a_{ave} \pm a_{std}$ %) ($a_{ave} \pm a_{std}$: average accuracy rate ± standard error over 20 repeats). The data sets was preprocessed by 1st derivative before modeling.

Data set	Type	PLS-DA	BSS	GA-PLS-DA	s-PLS-DA	EBSS
<i>Olive oil</i>	FTIR	94.2 ± 2.6	95.5 ± 3.8	94.6 ± 2.9	95.9 ± 2.5	95.6 ± 4.1
<i>Red wine</i>	FTIR	50.7 ± 12.9	53.2 ± 15.6	51.1 ± 14.9	50.0 ± 12	61.4 ± 12.9
<i>NIR Tablet</i>	NIR	86.8 ± 2.8	85.4 ± 3.0	85.5 ± 3.3	84.1 ± 2.6	85.6 ± 3.7

Raman Raman 80.8 ± 5.2 77.4 ± 7.2 74.9 ± 5.6 59.3 ± 6.3 51.7 ± 7.5

Tablet

Table S6. The number of selected variables ($n_{ave} \pm n_{std}$) ($n_{ave} \pm n_{std}$: average number of selected variable \pm standard error over 20 repeats). The data sets was preprocessed by 1st derivative before modeling.

Data set	Type	PLS-DA	BSS	GA-PLS-DA	s-PLS-DA	EBSS
<i>Olive oil</i>	FTIR	570	89 ± 65	59 ± 10	45 ± 22	24
<i>Red wine</i>	FTIR	842	49 ± 28	32 ± 11	47 ± 30	9
<i>NIR tablet</i>	NIR	404	39 ± 17	42 ± 8	56 ± 25	25
<i>Raman tablet</i>	Raman	3041	89 ± 65	59 ± 10	45 ± 22	44