

Supplementary Materials

List of Tables

Table S1. Demographic data in BC, BE, and HC.

| Characteristics | HC | BE | BC |
|------------------------------------|-----------|------------|------------|
| No. of subjects (n) | 68 | 56 | 53 |
| Age | | | |
| Mean ± SD | 34 ± 9.63 | 40 ± 11.92 | 52 ± 11.80 |
| Race n (%) | | | |
| Malay | 64 (94.1) | 51 (91.1) | 46 (86.8) |
| Non- malay | 4 (5.9) | 5 (8.9) | 7 (13.2) |
| Cancer stages n (%) | | | |
| Stage 0 | | | 2 (3.8) |
| Stage 1 | | | 5 (9.4) |
| Stage 2 | | | 19 (35.8) |
| Stage 3 | | | 12 (22.6) |
| Stage 4 | | | 15 (28.3) |
| Estrogen Receptor n (%) | | | |
| Positive | | | 35 (66.0) |
| Negative | | | 18 (34.0) |
| Progesterone Receptor n (%) | | | |
| Positive | | | 28 (52.8) |
| Negative | | | 25 (47.2) |
| HER status n (%) | | | |
| Positive | | | 18 (34.0) |
| Negative | | | 35 (66.0) |

Table S2. Important features identified by t-tests values, p-values (threshold <0.05), logarithmic p-values, and False Discovery Rate (FDR) parameters calculated for the most statistically significant compounds.

BC versus HC

| Metabolite | t.stat | p.value | -LOG10(p) | FDR |
|-------------------|--------|---------|-----------|----------|
| Kynurenate | 9.938 | <0.001* | 15.511 | <0.001* |
| Acetate | 8.5673 | <0.001* | 12.629 | <0.001* |
| 2-HBA | 7.6029 | <0.001* | 10.621 | <0.001* |
| Glucose | 7.1121 | <0.001* | 9.6169 | <0.001* |
| Pyruvate | 6.3532 | <0.001* | 8.1003 | <0.001* |
| Glycine | 5.8074 | <0.001* | 7.0466 | <0.001* |
| Tyrosine | 5.7917 | <0.001* | 7.0168 | <0.001* |
| Glycerol | 5.7232 | <0.001* | 6.8873 | <0.001* |
| Acetoacetate | 5.5223 | <0.001* | 6.5116 | <0.001* |
| Glutamine | 4.4 | <0.001* | 4.5363 | 0.000157 |
| Succinate | 4.3648 | <0.001* | 4.4783 | 0.000176 |
| Hydroxykynurenine | 4.1347 | <0.001* | 4.1058 | 0.000395 |
| Citrate | 4.1201 | <0.001* | 4.0825 | 0.000408 |

| | | | | |
|-------------------|---------|----------|--------|----------|
| Choline | 4.0557 | <0.001* | 3.9805 | 0.000488 |
| Acetone | 4.0539 | 0.000105 | 3.9777 | 0.000488 |
| Myo-inositol | 3.987 | 0.000134 | 3.8729 | 0.000602 |
| Phenylalanine.1 | 3.8403 | 0.000226 | 3.6468 | 0.000985 |
| Valine | 3.5848 | 0.000542 | 3.2659 | 0.002238 |
| Creatinine | 3.47 | 0.000794 | 3.1 | 0.003005 |
| Alanine | 3.126 | 0.002372 | 2.6249 | 0.008546 |
| Methylmalonate | 3.0399 | 0.003082 | 2.5112 | 0.0106 |
| Dimethylsulfone | 2.9513 | 0.004015 | 2.3964 | 0.013402 |
| GPC | -2.8725 | 0.005056 | 2.2962 | 0.016163 |
| Acetoin.1 | 2.8301 | 0.005714 | 2.243 | 0.018015 |
| Lactate | 2.7997 | 0.006232 | 2.2054 | 0.018951 |
| Malonate | 2.7869 | 0.006464 | 2.1895 | 0.019305 |
| Acetylglucosamine | 2.7292 | 0.007607 | 2.1188 | 0.022426 |
| Methylhistidine | -2.5993 | 0.010882 | 1.9633 | 0.030823 |

BE versus HC

| Metabolite | t.stat | p.value | -LOG10(p) | FDR |
|-----------------|---------|----------|-----------|----------|
| Acetone | -5.6561 | <0.001* | 6.7704 | <0.001* |
| Glucose | 4.8689 | <0.001* | 5.3392 | 0.00024 |
| Alanine | 4.7449 | <0.001* | 5.1236 | 0.000286 |
| Methylmalonate | -4.0598 | 0.000102 | 3.9901 | 0.001794 |
| Formate | -3.9794 | 0.000137 | 3.8639 | 0.002228 |
| Threonine | -3.9229 | 0.000167 | 3.7762 | 0.002457 |
| Glutamine | -3.8976 | 0.000183 | 3.7372 | 0.002457 |
| Acetate | 3.8738 | 0.000199 | 3.7007 | 0.002523 |
| Methylhistidine | -3.2958 | 0.00139 | 2.8569 | 0.013206 |
| Glycerol | -3.151 | 0.002189 | 2.6597 | 0.019967 |
| Phenylalanine | -3.0842 | 0.002687 | 2.5707 | 0.021126 |
| 3-HBA | -3.0619 | 0.002876 | 2.5412 | 0.021857 |
| Succinate | -3.0065 | 0.003398 | 2.4688 | 0.02499 |

BE versus BC

| Metabolite | t.stat | p.value | -LOG10(p) | FDR |
|---------------|---------|---------|-----------|---------|
| Glycerol | -9.9221 | <0.001* | 15.704 | <0.001* |
| Acetone | -9.3346 | <0.001* | 14.436 | <0.001* |
| 2-HBA | -9.1836 | <0.001* | 14.109 | <0.001* |
| Choline | -8.4984 | <0.001* | 12.635 | <0.001* |
| Tyrosine | -8.4727 | <0.001* | 12.58 | <0.001* |
| Threonine | -7.652 | <0.001* | 10.835 | <0.001* |
| Glutamine | -7.4254 | <0.001* | 10.36 | <0.001* |
| Glucose | -7.1436 | <0.001* | 9.7729 | <0.001* |
| Phenylalanine | -6.9239 | <0.001* | 9.32 | <0.001* |
| Succinate | -6.8707 | <0.001* | 9.2111 | <0.001* |
| Kynurenate | -6.2148 | <0.001* | 7.8901 | <0.001* |
| Citrate | -5.885 | <0.001* | 7.2449 | <0.001* |

| | | | | |
|-------------------|---------|----------|--------|----------|
| Dimethylsulfone | -5.7696 | <0.001* | 7.0227 | <0.001* |
| Acetoacetate | -5.2178 | <0.001* | 5.9881 | <0.001* |
| Myo-inositol | -5.0877 | <0.001* | 5.7516 | <0.001* |
| Glycine | -5.0261 | <0.001* | 5.6407 | <0.001* |
| Methylmalonate | -5.0024 | <0.001* | 5.5982 | <0.001* |
| Acetate | -4.8411 | <0.001* | 5.312 | <0.001* |
| Creatinine | -4.6888 | <0.001* | 5.0463 | <0.001* |
| 3-HBA | -4.4818 | <0.001* | 4.693 | <0.001* |
| Pyruvate | -4.4138 | <0.001* | 4.579 | 0.000107 |
| Hydroxykynurenine | -4.3515 | <0.001* | 4.4754 | 0.000134 |
| Acetylglucosamine | -4.336 | <0.001* | 4.4497 | 0.000137 |
| Malonate | -4.2269 | <0.001* | 4.2707 | 0.0002 |
| Creatine | -3.8584 | 0.000206 | 3.687 | 0.000721 |
| Valine | -3.3578 | 0.001124 | 2.9493 | 0.00366 |
| Lysine | 3.0862 | 0.002642 | 2.578 | 0.008253 |
| Formate | -2.5862 | 0.01119 | 1.9512 | 0.030372 |

*Strongly significant *p*-value <0.001

Table S3. Assignment of metabolites in BC, BE, and HC based on ¹H NMR spectroscopic data.

| No. | HMDB ID | Metabolite | ¹ H NMR characteristic signal(s) |
|-----|-----------|---------------------------|---|
| 1 | HMDB00001 | 1-Methylhistidine | 8.03 (s), 7.13 (s), 3.16 (m, overlap), 3.07 (m, overlap) |
| 2 | HMDB00008 | 2-Hydroxybutyrate (2-HBA) | 4.00 (m, overlap), 1.71 (m), 1.58 (m), 0.89 (t, overlap) |
| 3 | HMDB00042 | Acetate | 1.92 (s) |
| 4 | HMDB00060 | Acetoacetate | 2.25 (s) |
| 5 | HMDB00062 | L-Carnitine | 3.22 (s) |
| 6 | HMDB00064 | Creatine | 3.03 (s), 3.93 (s) |
| 7 | HMDB00094 | Citrate | 2.69 (d, J= 16.0 Hz), 2.53 (d, J= 16.0 Hz) |
| 8 | HMDB00097 | Choline | 3.17 (s) |
| 9 | HMDB00122 | D-Glucose | 5.23 (d, J=3.5 Hz), 4.64 (d, J= 8.0 Hz), 3.71-3.80 (m), 3.40-3.53 (m) |
| 10 | HMDB00123 | Glycine | 3.58 (s) |
| 11 | HMDB00131 | Glycerol | 3.65 (m), 3.58 (m) |
| 12 | HMDB00142 | Formate | 8.46 (s) |
| 13 | HMDB00148 | L-Glutamate | 2.33 (m), 2.12 (m, overlap), 2.04 (m, overlap) |
| 14 | HMDB00157 | Hypoxanthine | 8.20 (s), 8.18 (s) |
| 15 | HMDB00158 | L-Tyrosine | 7.19 (d, J= 8.0 Hz), 6.89 (d, J= 8.0 Hz) |
| 16 | HMDB00159 | L-Phenylalanine | 7.43 (m), 7.37 (m), 7.32 (m) |
| 17 | HMDB00161 | L-Alanine | 1.48 (d, J= 7.0 Hz) |
| 18 | HMDB00167 | L-Threonine | 4.25 (m) |
| 19 | HMDB00172 | L-Isoleucine | 1.01 (d, J=7.0 Hz), 0.92 (t, overlap) |
| 20 | HMDB00177 | L-Histidine | 7.84 (s), 7.08 (s) |
| 21 | HMDB00182 | Lysine | 3.02 (t, overlap), 1.92 (m), 1.73 (m), 1.50 (m) |
| 22 | HMDB00187 | L-Serine | 3.99 (m), 3.94 (m) |
| 23 | HMDB00190 | L-Lactate | 4.11 (m), 1.33 (d, J= 7.0 Hz) |
| 24 | HMDB00191 | L-Aspartate | 3.89 (m, overlap), 2.80 (dd, overlap), 2.70 (dd, J= 8.8, 17.2 Hz) |

| No. | HMDB ID | Metabolite | ¹ H NMR characteristic signal(s) |
|-----|-------------|--------------------------------------|--|
| 25 | HMDB00214 | Ornithine | 3.78 (m, overlap), 3.05 (t, J= 8.0 Hz), 1.94 (m), 1.82 (m), 1.73 (m, overlap) |
| 26 | HMDB00243 | Pyruvate | 2.38 (s) |
| 27 | HMDB00254 | Succinate | 2.40 (s) |
| 28 | HMDB00357 | 3-Hydroxybutyrate (3-HBA) | 4.14 (m), 2.29 (m, overlap), 1.17 (d, J= 6.8 Hz) |
| 29 | HMDB00562 | Creatinine | 3.04 (s), 4.05 (s) |
| 30 | HMDB00641 | L-Glutamine | 2.46 (m), 2.14 (m) |
| 31 | HMDB00687 | L-Leucine | 0.97 (d, overlap), 0.95 (d, overlap) |
| 32 | HMDB00691 | Malonate | 3.11 (s) |
| 33 | HMDB00696 | L-Methionine | 2.64 (t, overlap), 2.13 (s, overlap) |
| 34 | HMDB00883 | L-Valine | 1.04 (d, J= 7.0 Hz), 0.99 (d, J=7.0 Hz) |
| 35 | HMDB01659 | Acetone | 2.18 (s) |
| 36 | HMDB04983 | Dimethylsulfone | 3.15 (s) |
| 37 | HMDB0000211 | Myo-inositol | 4.06 (s), 3.60 (m, overlap), 3.53 (m, overlap), 3.32 (m) |
| 38 | HMDB0000086 | sn-Glycero-3-phosphocholine (GPC) | 3.22(s) |
| 39 | HMDB0000202 | Methylmalonate | 3.17 (m, overlap), 1.20 (d, J= 6.4 Hz) |
| 40 | HMDB000043 | Betaine | 3.89 (br s, overlap), 3.24 (s) |
| 41 | HMDB0000803 | N-Acetylglucosamine | 5.18 (br s), 2.08 (s) |
| 42 | HMDB0000715 | Kynurenate | 7.86 (m), 7.54 (m), 6.95 (m) |
| 43 | HMDB0000732 | 3-Hydroxykynurenine | 7.44 (m, overlap), 7.03 (br d), 6.73 (m) |
| 44 | HMDB0003243 | Acetoin | 2.23 (s), 1.37 (d, J=7.0 Hz) |

Table S4. Variable important to the projection (VIP) values of OPLS-DA model of BC versus BE versus HC groups.

| No | δ H | Metabolite | VIP value | CV-SE |
|----|------------|----------------|-----------|-------|
| 1 | 0.89 | Isoleucine | 3.31 | 0.68 |
| 2 | 1.33 | Lactate | 3.24 | 0.55 |
| 3 | 3.78 | Glucose | 3.21 | 0.29 |
| 4 | 3.86 | Glucose | 3.09 | 0.30 |
| 5 | 3.65 | Glycerol | 2.30 | 0.66 |
| 6 | 3.25 | Carnitine | 2.24 | 0.46 |
| 7 | 0.93 | Leucine | 2.13 | 0.29 |
| 8 | 5.26 | Glucose | 2.07 | 0.27 |
| 9 | 3.21 | GPC | 2.01 | 0.54 |
| 10 | 1.17 | 3-HBA | 1.96 | 0.54 |
| 11 | 4.66 | Glucose | 1.96 | 0.45 |
| 12 | 1.37 | Acetoin | 1.92 | 0.70 |
| 13 | 1.21 | Methylmalonate | 1.86 | 1.35 |
| 14 | 3.57 | Glycine | 1.77 | 0.23 |
| 15 | 2.13 | Glutamine | 1.73 | 0.81 |
| 16 | 0.97 | Valine | 1.62 | 0.28 |
| 17 | 1.49 | Alanine | 1.48 | 0.68 |
| 18 | 2.17 | Acetone | 1.47 | 0.55 |
| 19 | 2.05 | Glutamate | 1.26 | 0.80 |

| | | | | |
|----|------|--------------|------|------|
| 20 | 1.73 | 2-HBA | 1.23 | 0.28 |
| 21 | 3.05 | Creatinine | 1.09 | 0.48 |
| 22 | 2.25 | Acetoacetate | 0.98 | 0.57 |
| 23 | 1.05 | Valine | 0.92 | 0.26 |
| 24 | 2.41 | Succinate | 0.86 | 0.36 |
| 25 | 2.37 | Pyruvate | 0.86 | 0.45 |
| 26 | 4.14 | Lactate | 0.85 | 0.59 |
| 27 | 1.93 | Acetate | 0.84 | 0.18 |
| 28 | 3.01 | Creatine | 0.84 | 0.48 |
| 29 | 3.17 | Choline | 0.72 | 0.13 |

Table S5. Summary of the OPLS-DA model of BC versus BE versus HC groups.

| Component | R2X | R2X (cum) | Eigenvalue | R2 | R2Y (cum) | Q2 | Limit | Q2 (cum) | Eigenvalue Y | Significance |
|-----------------------|--------|-----------|------------|-------|-----------|-------|-------|----------|--------------|--------------|
| Model | | 0.765 | | | 0.654 | | | 0.563 | | |
| Predictive | | 0.246 | | | 0.654 | | | 0.563 | | |
| P1 | 0.202 | 0.202 | 29.1 | 0.36 | 0.36 | 0.337 | 0.01 | 0.337 | 1.51 | R1 |
| P2 | 0.0441 | 0.246 | 6.35 | 0.294 | 0.654 | 0.227 | 0.01 | 0.563 | 1.49 | R1 |
| Orthogonal in X(OPLS) | | 0.52 | | | 0 | | | | | |
| O1 | 0.196 | 0.196 | 28.2 | 0 | 0 | | | | | R1 |
| O2 | 0.206 | 0.402 | 29.7 | 0 | 0 | | | | | R1 |
| O3 | 0.0827 | 0.485 | 11.9 | 0 | 0 | | | | | R1 |
| O4 | 0.0345 | 0.52 | 4.97 | 0 | 0 | | | | | R1 |

Table S6. CV-ANOVA of the OPLS-DA model of BC versus BE versus HC groups.

| | SS | DF | MS | F | p | SD |
|--------------------|---------|-----|----------|--------|---|----------|
| Total corr. | 286 | 286 | 1 | | | 1 |
| Regression | 170.602 | 24 | 7.10843 | 16.139 | 0 | 2.66616 |
| Residual | 115.398 | 262 | 0.440449 | | | 0.663664 |

Table S7. Summary of the OPLS-DA model of cancer versus healthy control.

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2Y (cum) | Q2 | Limit | Q2 (cum) | Eigenvalue Y | Significance |
|-----------------------|-------|----------|------------|-------|-----------|------|-------|----------|--------------|--------------|
| Model | | 0.596 | | | 0.744 | | | 0.68 | | |
| Predictive | | 0.233 | | | 0.744 | | | 0.68 | | |
| P1 | 0.233 | 0.233 | 21.9 | 0.744 | 0.744 | 0.68 | 0.01 | 0.68 | 2 | R1 |
| Orthogonal in X(OPLS) | | 0.363 | | | 0 | | | | | |
| O1 | 0.202 | 0.202 | 19 | 0 | 0 | | | | | R1 |
| O2 | 0.161 | 0.363 | 15.1 | 0 | 0 | | | | | R1 |

Table S8. CV-ANOVA of the OPLS-DA model of BC versus HC.

| | SS | DF | MS | F | p | SD |
|--------------------|---------|----|----------|---------|--------------|---------|
| Total corr. | 93 | 93 | 1 | | | 1 |
| Regression | 63.2152 | 6 | 10.5359 | 30.7747 | 1.47566e-019 | 3.2459 |
| Residual | 29.7848 | 87 | 0.342354 | | | 0.58511 |

Table S9. Summary of the OPLS-DA model of BE versus HC.

| Component | R2X | R2X (cum) | Eigenvalue | R2 | R2Y (cum) | Q2 | Limit | Q2 (cum) | Eigenvalue Y | Significance |
|-----------------------|--------|-----------|------------|-------|-----------|-------|-------|----------|--------------|--------------|
| Model | | 0.406 | | | 0.693 | | | 0.596 | | |
| Predictive | | 0.0805 | | | 0.693 | | | 0.596 | | |
| P1 | 0.0805 | 0.0805 | 7.64 | 0.693 | 0.693 | 0.596 | 0.01 | 0.596 | 2 | R1 |
| Orthogonal in X(OPLS) | | 0.325 | | | 0 | | | | | |
| O1 | 0.234 | 0.234 | 22.2 | 0 | 0 | | | | | R1 |
| O2 | 0.0916 | 0.325 | 8.7 | 0 | 0 | | | | | R1 |

Table S10. CV-ANOVA of the OPLS-DA model of BE versus HC.

| | SS | DF | MS | F | p | SD |
|--------------------|---------|----|----------|---------|--------------|----------|
| Total corr. | 94 | 94 | 1 | | | 1 |
| Regression | 55.9841 | 6 | 9.33068 | 21.5989 | 1.90004e-015 | 3.05462 |
| Residual | 38.0159 | 88 | 0.431999 | | | 0.657266 |

Table S11. Summary of the OPLS-DA model of BC versus BE.

| Component | R2X | R2X (cum) | Eigenvalue | R2 | R2Y (cum) | Q2 | Limit | Q2 (cum) | Eigenvalue Y | Significance |
|-----------------------|-------|-----------|------------|-------|-----------|-------|-------|----------|--------------|--------------|
| Model | | 0.552 | | | 0.768 | | | 0.694 | | |
| Predictive | | 0.189 | | | 0.768 | | | 0.694 | | |
| P1 | 0.189 | 0.189 | 18.7 | 0.768 | 0.768 | 0.694 | 0.01 | 0.694 | 2 | R1 |
| Orthogonal in X(OPLS) | | 0.363 | | | 0 | | | | | |
| O1 | 0.22 | 0.22 | 21.7 | 0 | 0 | | | | | R1 |
| O2 | 0.143 | 0.363 | 14.2 | 0 | 0 | | | | | R1 |

Table S12. CV-ANOVA of the OPLS-DA model of BC versus BE.

| | SS | DF | MS | F | p | SD |
|--------------------|---------|----|----------|---------|-------------|----------|
| Total corr. | 98 | 98 | 1 | | | 1 |
| Regression | 68.0567 | 6 | 11.3428 | 34.8504 | 1.1406e-021 | 3.3679 |
| Residual | 29.9433 | 92 | 0.325471 | | | 0.570501 |

Table S13. All matched pathways according to the p values (threshold <0.05) from the pathway enrichment analysis and pathway impact values.

| Pathway | Total | Expected | Hits | p value | FDR | Impact |
|---|-------|----------|------|-----------|----------|--------|
| Aminoacyl-tRNA biosynthesis | 48 | 1.363 | 15 | <0.001* | <0.001* | 0.167 |
| Glyoxylate and dicarboxylate metabolism | 32 | 0.908 | 8 | <0.001* | <0.001* | 0.180 |
| Alanine, aspartate and glutamate metabolism | 28 | 0.795 | 7 | <0.001* | 0.000193 | 0.534 |
| Glycine, serine and threonine metabolism | 33 | 0.937 | 7 | <0.001* | 0.00047 | 0.513 |
| Valine, leucine and isoleucine biosynthesis | 8 | 0.227 | 4 | <0.001* | 0.000613 | 0.000 |
| Arginine biosynthesis | 14 | 0.397 | 4 | 0.0004603 | 0.006444 | 0.178 |
| Butanoate metabolism | 15 | 0.426 | 4 | 0.0006148 | 0.007377 | 0.111 |
| Histidine metabolism | 16 | 0.454 | 4 | 0.0008028 | 0.008429 | 0.221 |
| Phenylalanine, tyrosine and tryptophan biosynthesis | 4 | 0.114 | 2 | 0.0045588 | 0.038557 | 1.000 |
| Valine, leucine and isoleucine degradation | 40 | 1.136 | 5 | 0.0045901 | 0.038557 | 0.023 |
| Glycolysis / Gluconeogenesis | 26 | 0.738 | 4 | 0.0053571 | 0.040908 | 0.130 |
| Synthesis and degradation of ketone bodies | 5 | 0.142 | 2 | 0.007461 | 0.052227 | 0.600 |
| Nitrogen metabolism | 6 | 0.170 | 2 | 0.01099 | 0.065942 | 0.000 |
| D-Glutamine and D-glutamate metabolism | 6 | 0.170 | 2 | 0.01099 | 0.065942 | 0.500 |
| Citrate cycle (TCA cycle) | 20 | 0.568 | 3 | 0.017368 | 0.09726 | 0.169 |
| Arginine and proline metabolism | 38 | 1.079 | 4 | 0.02064 | 0.10836 | 0.209 |
| Pyruvate metabolism | 22 | 0.625 | 3 | 0.022553 | 0.11144 | 0.267 |
| Phenylalanine metabolism | 10 | 0.284 | 2 | 0.030672 | 0.14314 | 0.357 |
| Glutathione metabolism | 28 | 0.795 | 3 | 0.042636 | 0.18849 | 0.108 |

*Strongly significant *p*-value <0.001

List of Figures

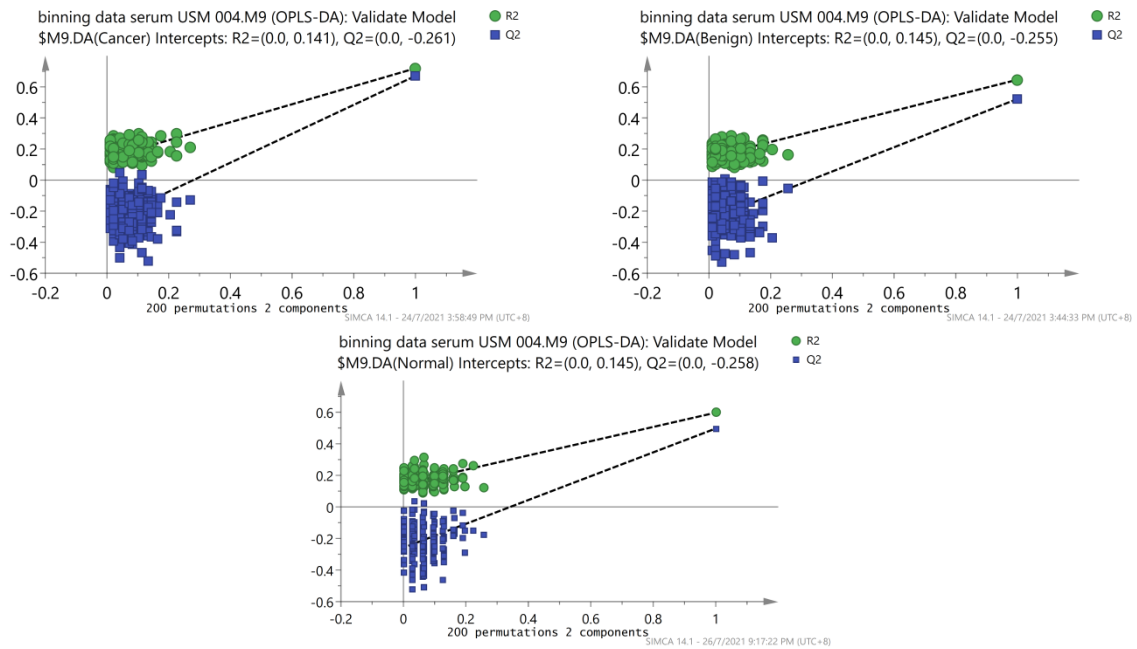


Figure S1. Permutation test results of the OPLS-DA model of BC versus BE versus HC groups.

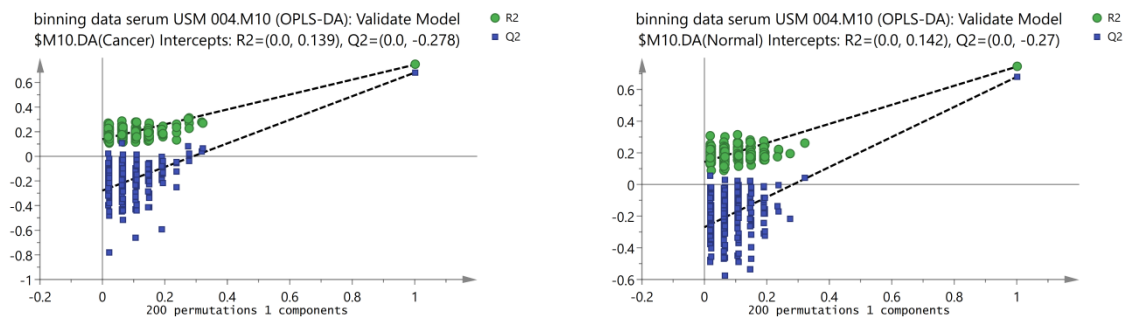


Figure S2. Permutation test results of the OPLS-DA model of BC versus HC.

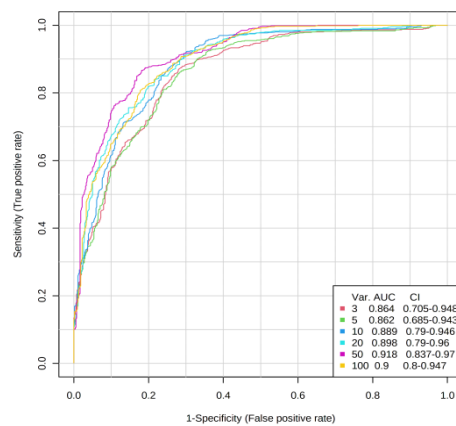


Figure S3. ROC curves based on the cross validation (CV) performance.

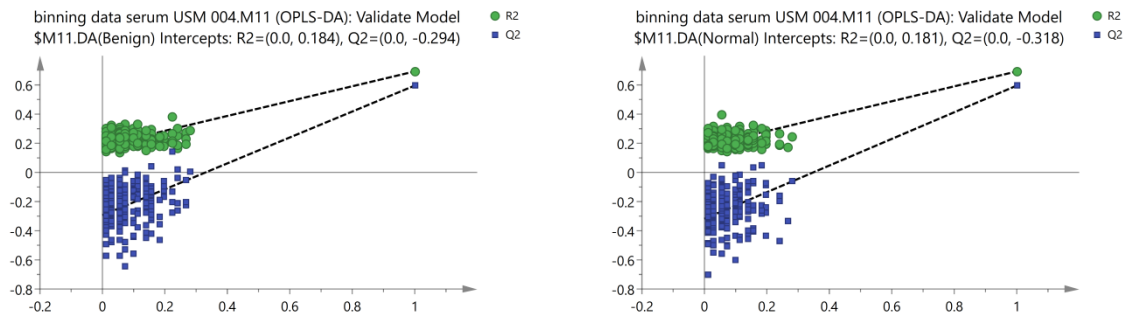


Figure S4. Permutation test results of the OPLS-DA model of BE versus HC.

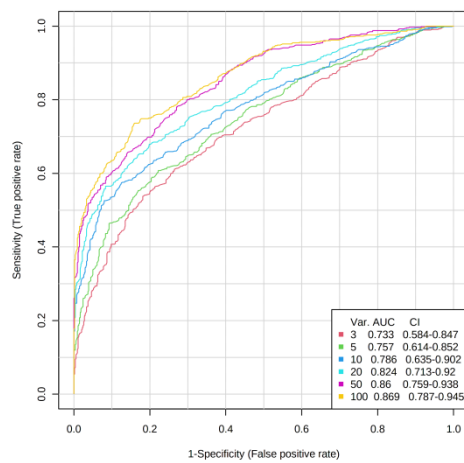


Figure S5. ROC curves based on the cross validation (CV) performance.

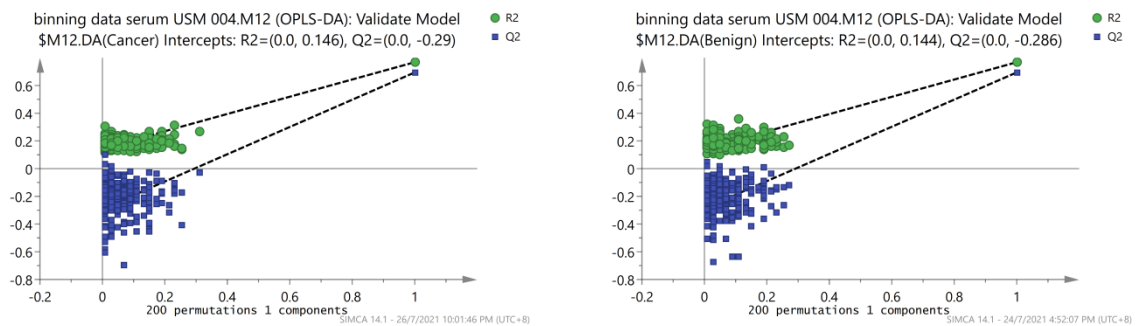


Figure S6. Permutation test results of the OPLS-DA model of BC versus BE.

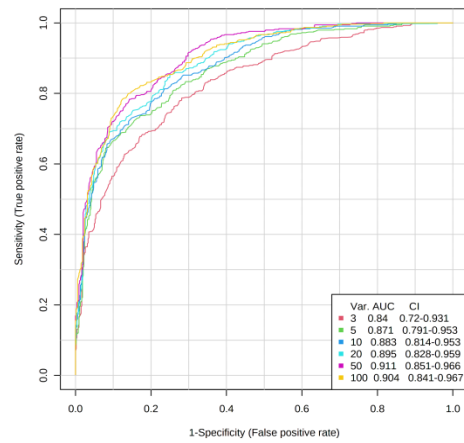


Figure S7. ROC curves based on the cross validation (CV) performance.

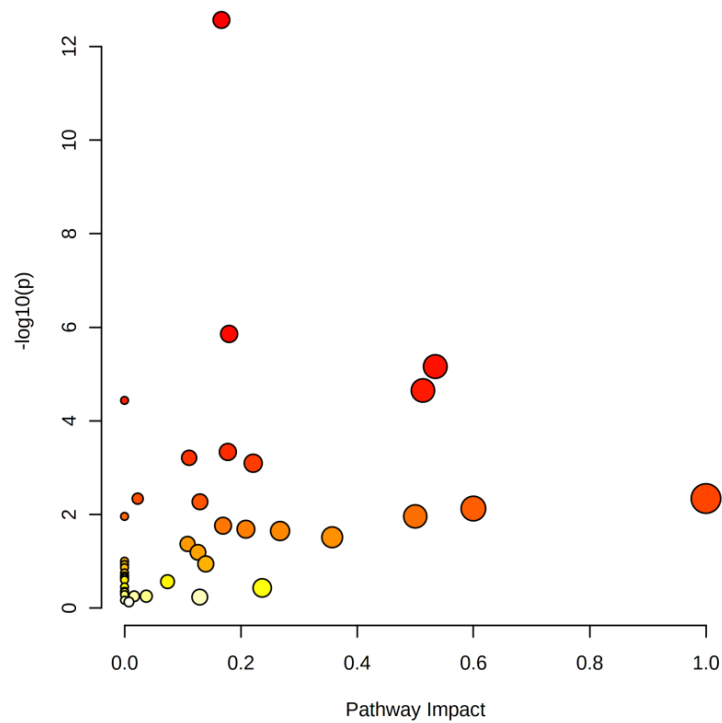


Figure S8. Overview of pathway analysis.