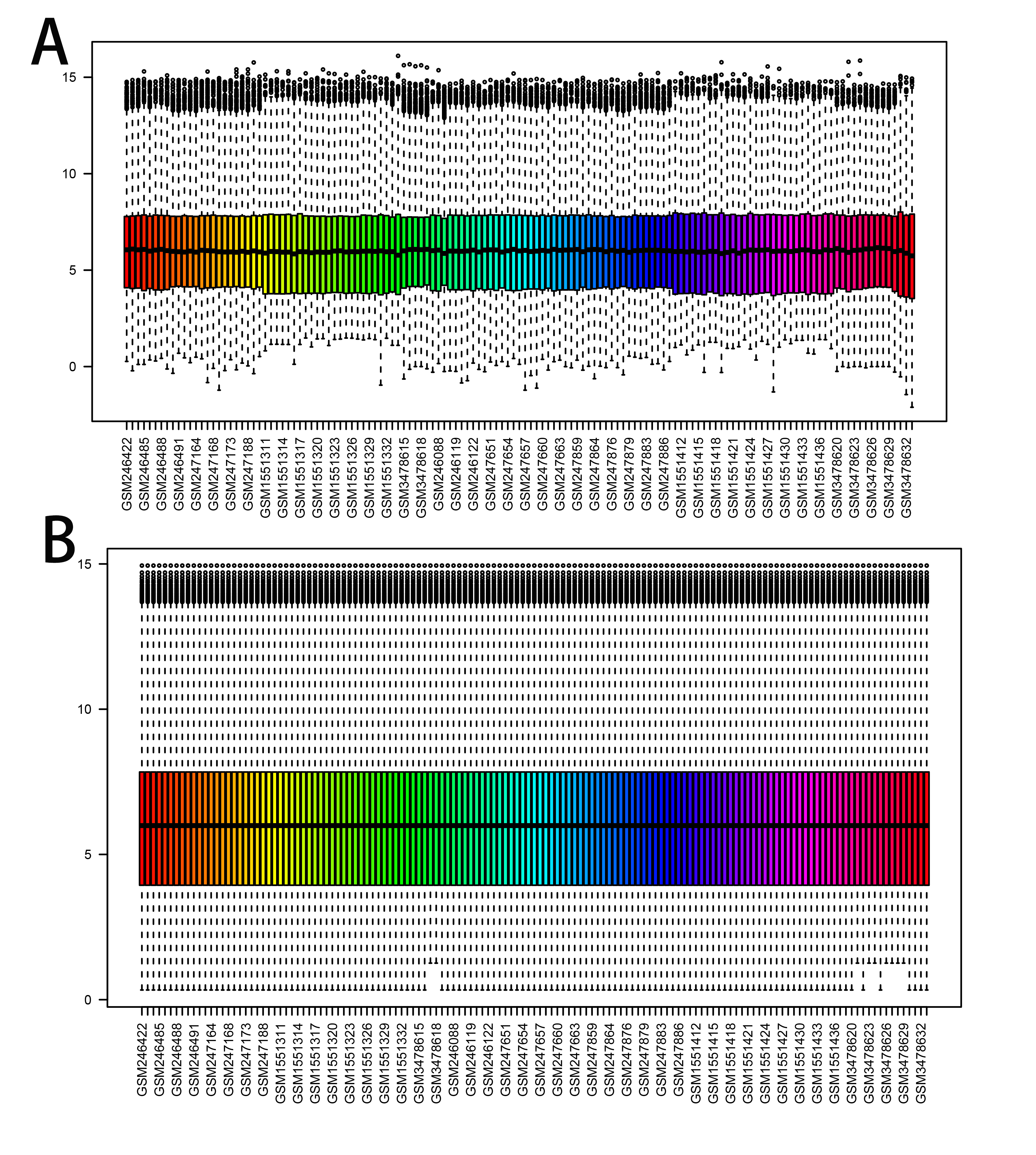
Supplementary material



Supplementary Fig. 1. Comparison of batch effect removal processes. (A) Before the batch effect is removed. (B) After the batch effect is removed.

Supplementary Table 1. Primer sequences.

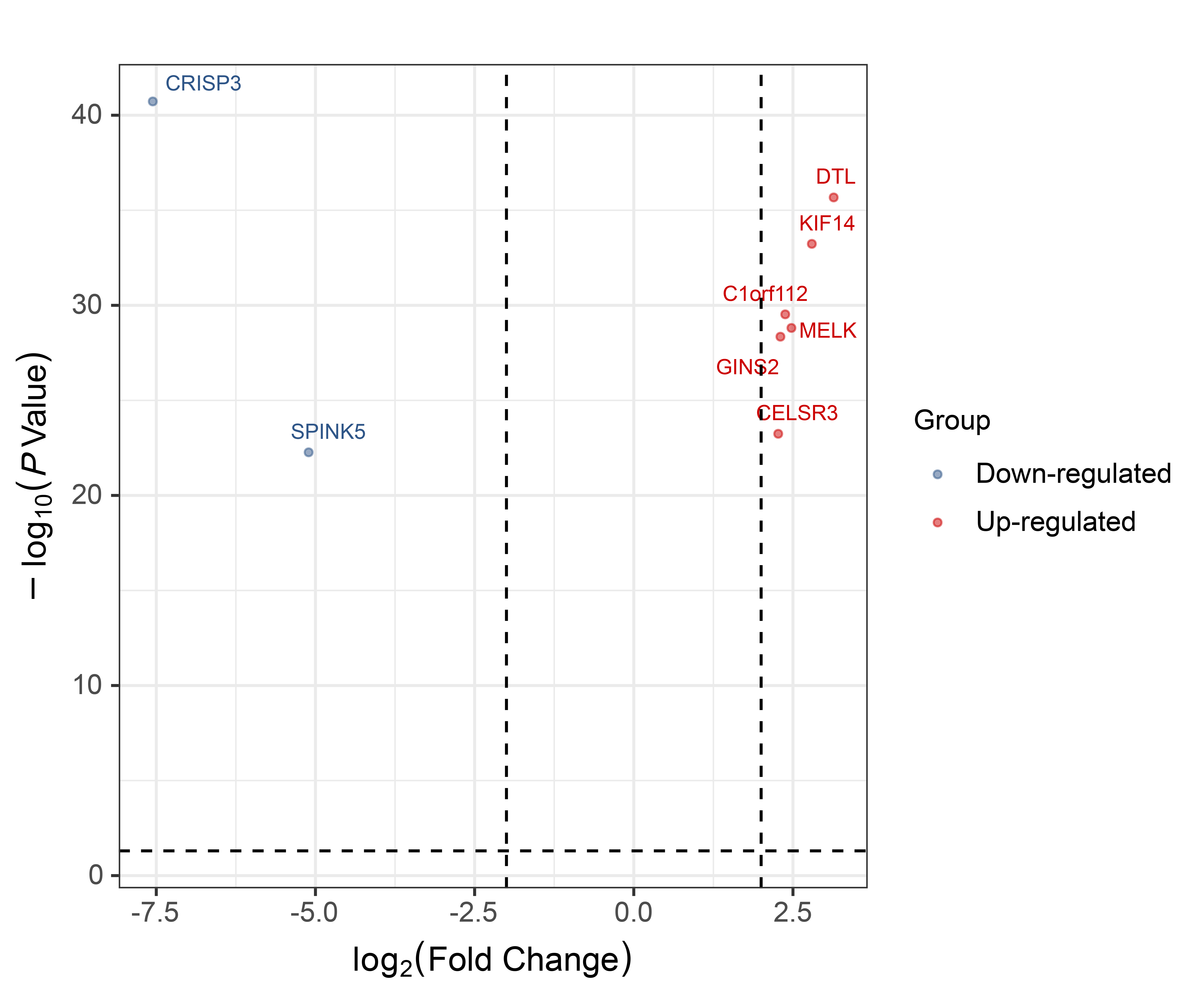
|  |  |  |
| --- | --- | --- |
| Gene | Forward Primer Sequence (5′-3′) | Reverse Primer Sequence (5′-3′) |
| *MELK* | GCTGCAAGGTATAATTGATGGA | CAGTAACATAATGACAGATGGGC |
| *CRISP3* | ACCTTGTGCCAGTTGCCCAGA | CCCTGACCAACTGATGTTTACAGG |
| *GINS2* | CAGAAATGTCGCCTGCTCC | GGATTTCGTCTGCCTTCG |
| *DTL* | ACCAGGAAGACCTTAGTA | AAGCATACGGACTGATAG |
| *C1orf112* | ATGTCTGCTGAAGTCATCC | ATCCTCACCAATCTCCTGCC |
| *KIF14* | TTCAGAACACCTCTGCAGGA | ACTCATGAAGACTACCTGGG |
| *SPINK5* | TGCTTTACGTGGCAAGGAG | GTAAGAACCTTTGTGATTACATGGG |
| *CELSR3* | CCTGCCAGCCAGGTTACTAC | ATCCTGTGCTCACAGTGGTG |

Supplementary Table 2. The detail information of DEGs’ logFC.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| DEGs | LogFC | DEGs | LogFC | DEGs | LogFC | DEGs | LogFC |
| CDKN2A | 3.951059602 | PDGFD | −2.168890498 | TM7SF2 | −2.39775579 | ID4 | −2.070803648 |
| CRISP3 | −7.555016662 | SCNN1B | −3.314723305 | SERPINB1 | −2.117680502 | TGM1 | −3.469848378 |
| DTL | 3.139390004 | ASPM | 2.310217782 | PAMR1 | −3.095164228 | ZNF750 | −3.060536197 |
| MCM2 | 2.705832977 | C2orf54 | −2.523304542 | CLCA4 | −4.688891951 | PLA2G7 | 2.240085728 |
| KNTC1 | 2.392475511 | KLK13 | −3.254633807 | NETO2 | 2.176448193 | COL21A1 | −2.642337691 |
| KIF14 | 2.796347019 | TGM3 | −4.166596661 | ISG15 | 2.394981235 | MALL | −2.484944903 |
| CRNN | −7.968868647 | HPGD | −3.621081746 | ALOX12B | −4.463344813 | LYPD3 | −2.872839421 |
| SLURP1 | −5.257465222 | KLK11 | −4.1348097 | ZNF185 | −2.087172108 | IFI44L | 3.276432524 |
| RFC4 | 2.514775021 | SULT2B1 | −2.840640297 | OIP5 | 2.070464433 | SPRR1B | −3.674044777 |
| C1orf112 | 2.378705059 | TPX2 | 2.078705463 | PRSS2 | −2.101543174 | FCGBP | −3.145094373 |
| RAD54L | 2.581582629 | MCM10 | 2.264514923 | SLC24A3 | −2.788876463 | RSAD2 | 2.020980534 |
| UPK1A | −5.218432239 | KRT1 | −6.448235245 | LHX2 | 2.132419451 | AKR1B10 | −3.216816555 |
| KIF4A | 2.746025284 | GLTP | −2.046695688 | KLK8 | −2.11676144 | KRT13 | −4.481577161 |
| MELK | 2.475723926 | ACOX2 | −2.386280253 | CYP3A5 | −2.464689327 | MMP1 | 3.258351507 |
| GINS2 | 2.302989093 | TRIP13 | 2.135151718 | EREG | −2.921756468 | P2RY2 | −2.123213588 |
| ATAD2 | 2.335332672 | KIFC1 | 2.079830177 | MOCOS | 2.00784053 | DEFB1 | −2.396690557 |
| SLC5A1 | −2.498600663 | IVL | −4.028401126 | HOXC6 | 2.885043819 | PTGDS | −2.290808126 |
| POLQ | 2.08338384 | APOD | −3.420243416 | MLF1 | 2.285735746 | MMP9 | 2.338384546 |
| MAL | −5.963530319 | FCER1A | −2.639241736 | SPINK2 | −2.076395417 | CXCL13 | 2.667378252 |
| SPAG5 | 2.477285201 | DSG1 | −5.283346991 | KRT4 | −4.893917014 | FMO2 | −2.588510753 |
| EDN3 | −3.801283266 | CENPE | 2.588038683 | RHCG | −5.005401665 | LCN2 | −2.657182692 |
| THSD4 | −2.548403082 | GINS1 | 2.025252663 | CYP2C18 | −2.958105154 | CALML3 | −2.39011135 |
| NEK2 | 2.131115125 | CRCT1 | −5.352416527 | CXCL14 | −3.149159707 | PI3 | −2.868823994 |
| FANCI | 2.133422274 | KRT2 | −3.686407996 | PLOD2 | 2.048446039 | RAPGEFL1 | −2.070795504 |
| BBOX1 | −4.561651156 | TGM5 | −2.826969687 | ALOX15B | −2.198065675 | FGFBP1 | −2.402350254 |
| WDHD1 | 2.045785732 | CA9 | 2.974157688 | KLK10 | −3.253530795 | DCN | −2.036978399 |
| HELLS | 2.655662896 | PPL | −2.346130858 | RIBC2 | 2.015042067 | SERPINB13 | −2.381615539 |
| PHYHIP | −2.702137663 | CLIC3 | −4.306964506 | BEX4 | −2.54027407 | KRT7 | 2.05387885 |
| PRC1 | 2.121250472 | ZBED2 | −3.659770463 | LOR | −3.475544207 | ISL1 | −2.095678602 |
| NUP210 | 2.047251937 | CEP55 | 2.413457947 | AIM2 | 3.105564636 | SEPP1 | −2.285613565 |
| AURKA | 2.361915155 | KIF23 | 2.114744842 | OCA2 | −2.145458027 | SPARCL1 | −2.185165907 |
| ESR1 | −2.071043639 | SCEL | −3.884541279 | IL18 | −2.47387962 | FLG | −2.935568689 |
| DNA2 | 2.223120964 | USP18 | 2.312682124 | SLC39A2 | −2.155645809 | CEL | 2.283359926 |
| CHAF1B | 2.364391295 | KCNK7 | −2.131210405 | NOD2 | −2.743984361 | NEFH | 2.188179884 |
| KIF2C | 2.180304595 | SLC27A6 | −2.785202508 | CDKN3 | 2.036729694 | WISP2 | −2.299643571 |
| EXO1 | 2.21724361 | NUSAP1 | 2.072077541 | ECM1 | −2.378027933 | COL17A1 | −2.102034573 |
| FOXM1 | 2.890731886 | CRYAB | −3.740246267 | TMPRSS11E | −3.722452914 | SERPINB2 | −2.41406842 |
| ALOX12 | −5.337213585 | TP53AIP1 | −2.193285724 | KLK6 | −2.930756299 | CXCL11 | 2.128064142 |
| TOP2A | 2.28004821 | FOXD1 | 2.656747601 | C1orf116 | −2.417626883 | IDO1 | 2.021920541 |
| CDC7 | 2.046534461 | NDRG4 | −3.106381778 | SPRR2B | −4.512100016 | CXCL9 | 2.376002594 |
| KLK12 | −5.29701447 | ZNF415 | −2.351063296 | KLK7 | −3.236725004 | RAB25 | −2.023292365 |
| CENPN | 2.051139607 | KLF4 | −2.467795232 | TMPRSS11D | −3.678889044 | S100A7 | −3.349737472 |
| CFD | −3.033499613 | RRM2 | 2.218951861 | SPRR2C | −3.914021988 | SERPINB3 | −2.400365369 |
| ACPP | −2.954453742 | KIF20A | 2.223008689 | CD207 | −2.164959494 | CLCA2 | −2.333240369 |
| SYCP2 | 3.136093777 | GSTA4 | −2.276840162 | EVPL | −2.205960146 | GLDC | 2.017957635 |
| PPP1R3C | −3.805236262 | TTK | 2.018795674 | CCL18 | 2.131174013 | S100A9 | −2.425751975 |
| CELSR3 | 2.268414601 | CRABP2 | −3.266501367 | NDN | −2.385275627 | KRT6B | −2.27687984 |
| NDC80 | 2.125232414 | CD177 | −2.258052568 | LY6G6C | −2.3887005 | AREG | −2.007423554 |
| SOSTDC1 | −4.243958267 | S100A12 | −3.680498891 | CEACAM7 | −2.628877181 | SCGB2A2 | −2.102611044 |
| HOPX | −3.673729213 | PRSS3 | −3.525716529 | ENO2 | 2.186206366 | LY6D | −2.626263607 |
| RAD51AP1 | 2.083552387 | SNX10 | 2.27166467 | MMP12 | 3.332124445 | DSG3 | −2.053416567 |
| AR | −2.300632483 | GDPD3 | −2.263774824 | PTK6 | −2.211100841 | NTS | 2.436089575 |
| EMP1 | −2.52054047 | EPB41L3 | −2.676671764 | CLDN17 | −2.193100025 | TCN1 | −2.128047851 |
| CDC6 | 2.24587073 | HSPB8 | −2.409892152 | SPRR3 | −4.806920788 | SERPINB4 | −2.60517458 |
| SPINK5 | −5.106939172 | DLGAP5 | 2.152085262 | DSC2 | −2.63413576 | KRT15 | −2.27677172 |
| NCAPH | 2.089031666 | APOC1 | 2.81417788 | PLA2G3 | −2.162307691 | CALML5 | −2.16488841 |
| CDCA8 | 2.480173348 | IL1R2 | −3.123476445 | EPHX2 | −2.079388047 | TFF3 | −2.105331664 |
| SYNGR3 | 3.319518439 | CDA | −3.460553618 | CST6 | −2.895349872 | SCGB2A1 | −2.006747398 |
| MCM5 | 2.193699912 | CRISP2 | −2.761440491 | SPRR1A | −3.935072951 | KRT14 | −2.0963145 |
| GPX3 | −3.720114954 | KIF15 | 2.081954617 | SPP1 | 2.815627393 | OLFM4 | −2.056288183 |
| GYS2 | −3.530231285 | - | - | - | - | - | - |

Supplementary Table 3. The expression of eight disease-specific genes (mean ± sd).

|  |  |  |
| --- | --- | --- |
|  | HCerEpC | Hela |
| MELK | 1.021 ± 0.250 | 1.453 ± 0.091 |
| CRISP3 | 1.014 ± 0.207 | 0.286 ± 0.052 |
| GINS2 | 1.074 ± 0.515 | 3.194 ± 0.209 |
| DLT | 1.019 ± 0.248 | 1.749 ± 0.211 |
| C1orf112 | 1.002 ± 0.054 | 4.014 ± 0.722 |
| KIF14 | 1.003 ± 0.123 | 3.273 ± 0.751 |
| SPINK5 | 1.027 ± 0.305 | 0.367 ± 0.069 |
| CELSR3 | 1.004 ± 0.116 | 2.649 ± 0.357 |



Supplementary Fig. 2. Volcanic map of eight disease-specific genes.