Discovery of Signature Cyclic Immonium Ion for Lactyllysine Reveals Widespread and Functional Lactylation on Non-histone Proteins in Human

Ning Wan[1, †], Nian Wang[1, †], Ying Kong[2], Xinmiao Wang[1], Dexiang Wang[1], Siqin Yu[1], Wenjie Lu[2], Yinxue Zhu[2], Chenxi Yang[3], Chang Shao[2], Runbin Sun[1], Nanxi Wang[4], Haiping Hao[1, 2, *], Hui Ye[1, *]

[1] Jiangsu Provincial Key Laboratory of Drug Metabolism and Pharmacokinetics, State Key Laboratory of Natural Medicines, China Pharmaceutical University, Tongjiaxiang No. 24, Nanjing 210009, Jiangsu, China,
[2] School of Pharmacy, China Pharmaceutical University, Tongjiaxiang No. 24, Nanjing 210009, Jiangsu, China,
[3] School of Biological Science and Medicinal Engineering, Southeast University, Sipailou No. 2, Nanjing 210096, Jiangsu, China,
[4] School of Pharmacy, Nanjing University of Chinese Medicine, Xianlindadao No. 138, Nanjing 210023, Jiangsu, China.

[†] These authors contributed equally to this manuscript.

* Correspondence: haipinghao@cpu.edu.cn (H.H.), Tel: 86-25-83271179,
cpuyehui@cpu.edu.cn (H.Y.), Tel: 86-25-83271191.

Figure S1

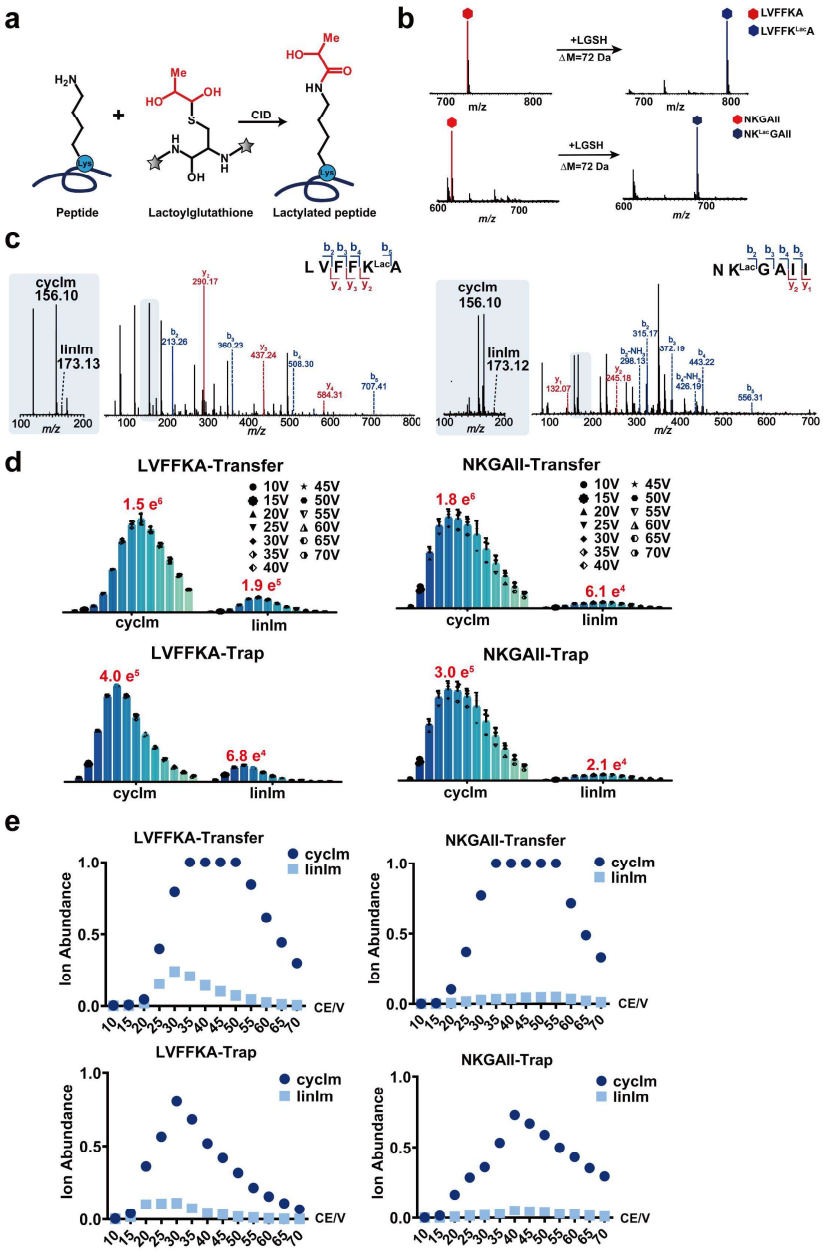

Figure S1. Validation of the production of a signature cyclic immonium ion from lysine lactylated peptides.

(a) Illustration of lysine lactylation introduced by a nonenzymatic acyl transfer reaction with LGSH. DCC, Dicyclohexylcarbodiimide; HOBT, 1-Hydroxybenzotriazole. (b) Mass spectra of the successfully lactylated model peptides. (c) MS/MS spectra of the lactylated model peptides. (d-e) Collisional energy (CE) and collision cell (trap and transfer cell on a SYNAPT G2 Si Q/TOF) influenced the ion counts (d) and relative ion abundance (e) of the cyclic immonium ion (cyclm) and the linear immonium ion (linlm) in MS/MS spectra. Relative ion abundance is normalized against the ion count of the base peak among the matched ions in examined MS/MS spectrum.

Figure S2

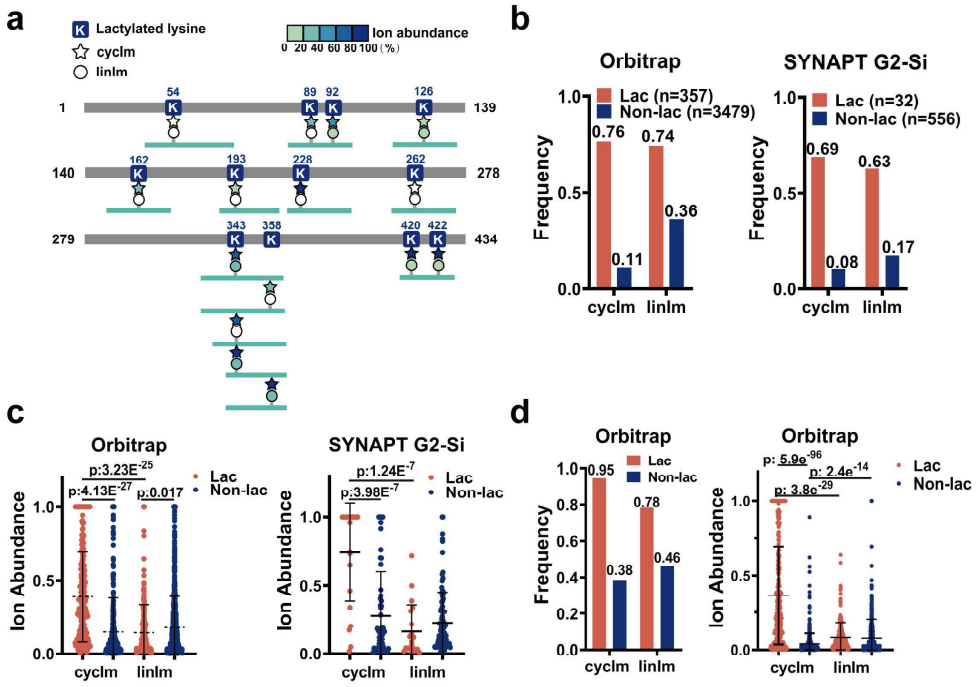

Figure S2. Production efficiency of the signature cyclm ions from lysine lactylation peptides is verified using different instruments and search engines.

(a) Overview of the assigned lactylation sites of *in vitro* lactylated recombinant ENO1. The immonium ion abundances in MS/MS spectra of lactylated peptides detected on a SYNAPT G2-Si Q/TOF were annotated. (b) Production frequency of the cyclm and linlm ions in lactylated and non-lactylated peptide-spectrum matches (PSMs) of ENO1-digested peptides using data obtained on a SYNAPT G2-Si Q/TOF and an Orbitrap Fusion Lumos, respectively. (c) Relative ion abundance of the cyclm and linlm ions in PSMs shown in (b). Only cyclm and linlm ions of non-zero ion counts were summarized for comparison. Data represent mean±SD and statistical significance was determined using the unpaired two-tailed t-test. (d) Frequency and ion abundance of the cyclm and linlm ions in LGSH-mediated lactylated and non-lactylated PSMs of recombinant ENO1-digested peptides searched by the Proteome Discoverer (PD) software.

Figure S3

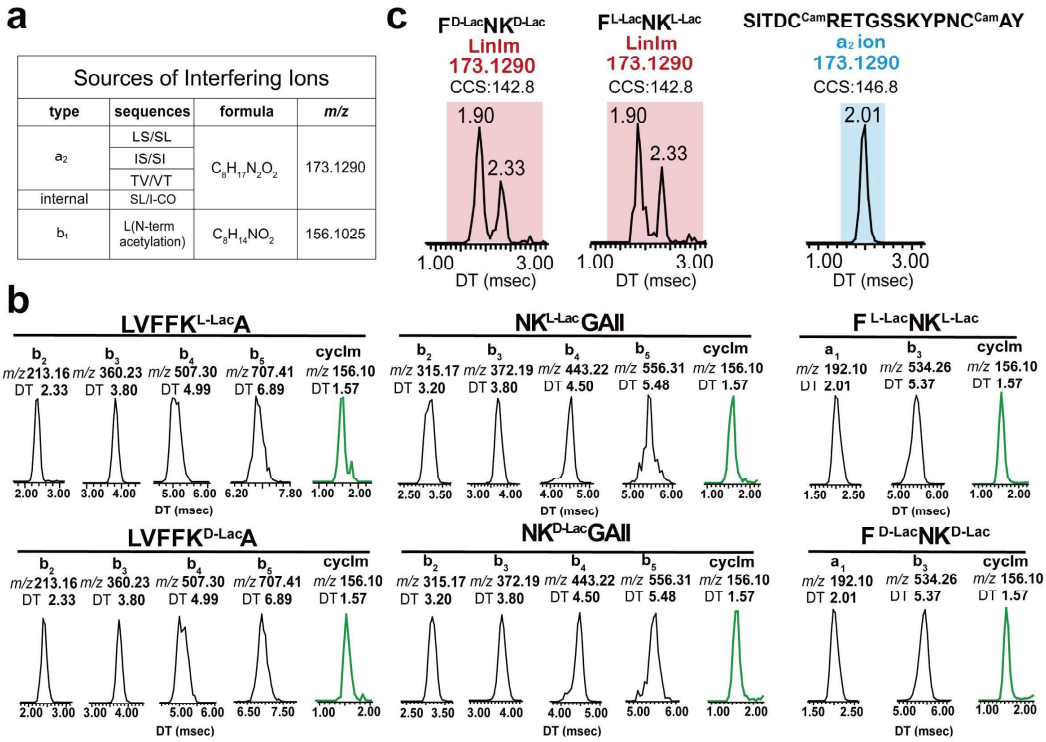

Figure S3. Ion mobility mass spectrometry probes the gas-phase behaviors of the cyclm and linlm ions of lactyllysine.

(a) Sources of interfering MS/MS ions of the cyclm and linlm ions of lactyllysine. (b) Drift time of MS/MS fragmentation-produced a-, b-type ions and cyclm ion generated by L- and D-lactylated model peptides. The collision cross section (CCS) value of the cyclm ion is measured as 131.1 Å. (c) Arrival time distributions (ATDs) resolve the unspecific linlm ion of lactyllysine from an isobaric, interfering a_2 ion produced from an RNase A-digested peptide. Cam, carbamidomethylation.

Figure S4

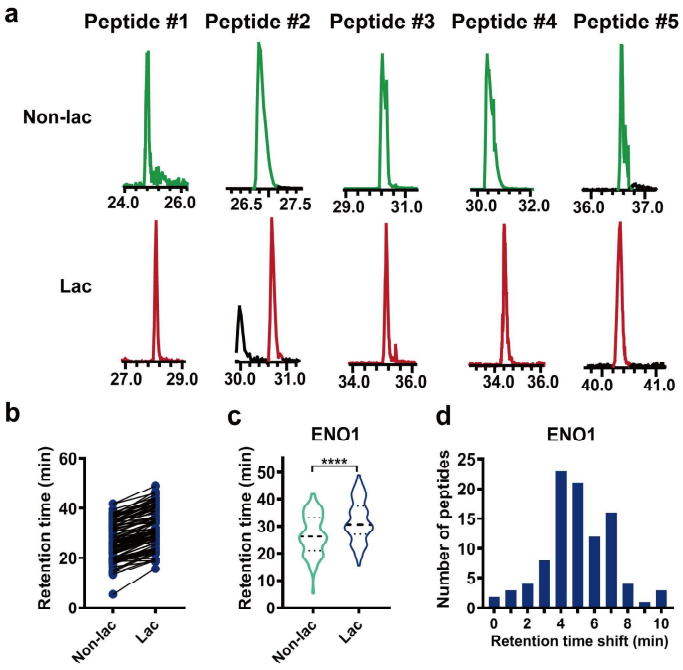

Figure S4. Chromatographic behaviors of lactylated *vs.* non-lactylated peptides detected on a nanoLC-Orbitrap MS.

(a) Pairwise comparison of extracted ion chromatograms (XICs) of the lactylated and non-lactylated peptides numbered Peak 6-13 (refer to the *in vitro* lactylated peptide library in **Table S1**) using data obtained on a SYNAPT G2-Si Q/TOF. (b) Retention time of lactylated and non-lactylated peptide pairs from recombinant ENO1 using data obtained on an Orbitrap. (c) Retention time distribution of lactylated and non-lactylated peptide pairs using data obtained on an Orbitrap. (d) Retention time shifts following lactylation using ENO1-digested peptides detected on an Orbitrap.

Figure S5

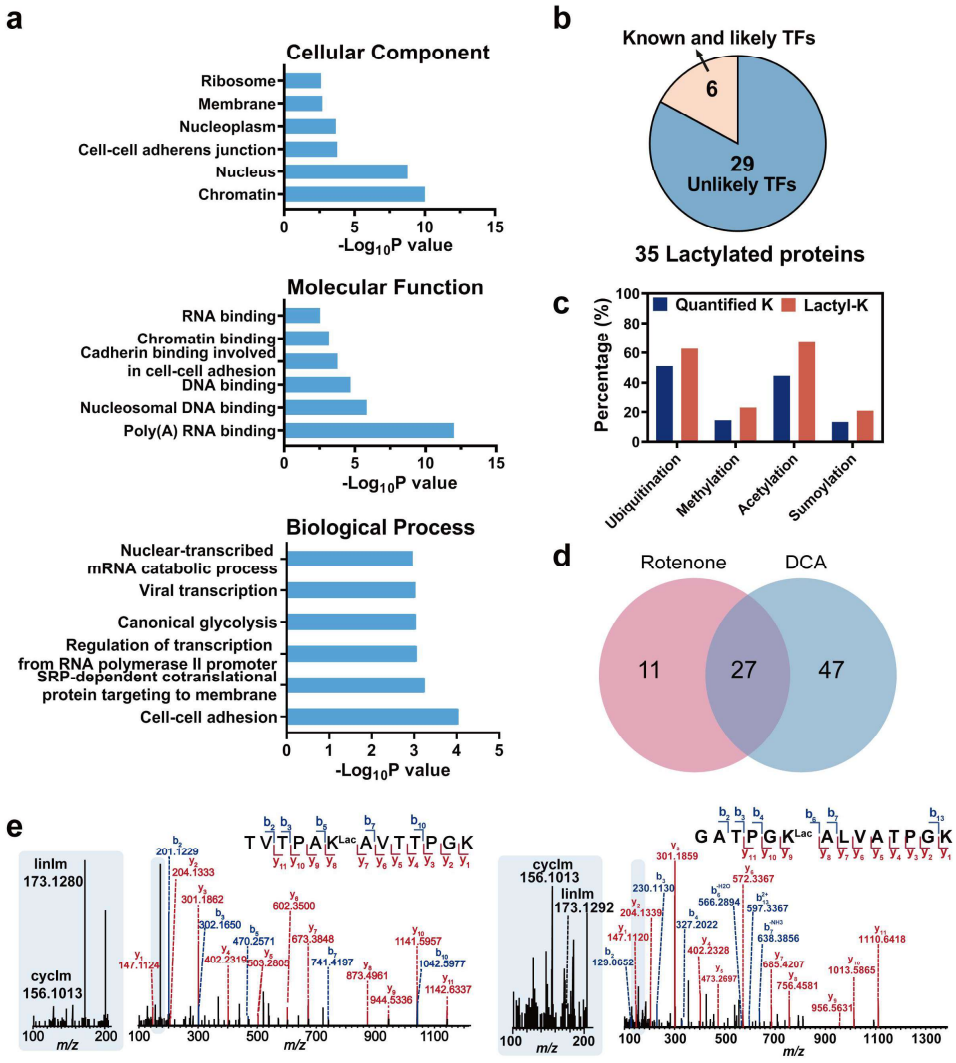

Figure S5. Functional and quantitative analysis of lactylated proteins detected in SILAC-processed, affinity-enriched MCF-7 cell proteome in response to rotenone and DCA treatment.

(a) Cellular component (CC), molecular function (MF), biological processes (BP) analysis of the identified lactylated non-histone proteins. (b) Transcription factor (TF) analysis of the identified lactylated non-histone proteins (n=35). (c) Post-translational modification (PTM) category analysis of all quantified lysines *vs.* lactylated lysines according to the examined SILAC proteomics data. (d) Venn diagram shows the overlap of the quantified lactylated peptides in response to rotenone- and DCA-treatment. (e) Representative MS/MS spectra of nucleolin (NCL) peptides carrying lactylation at K102 and K116.

Figure S6

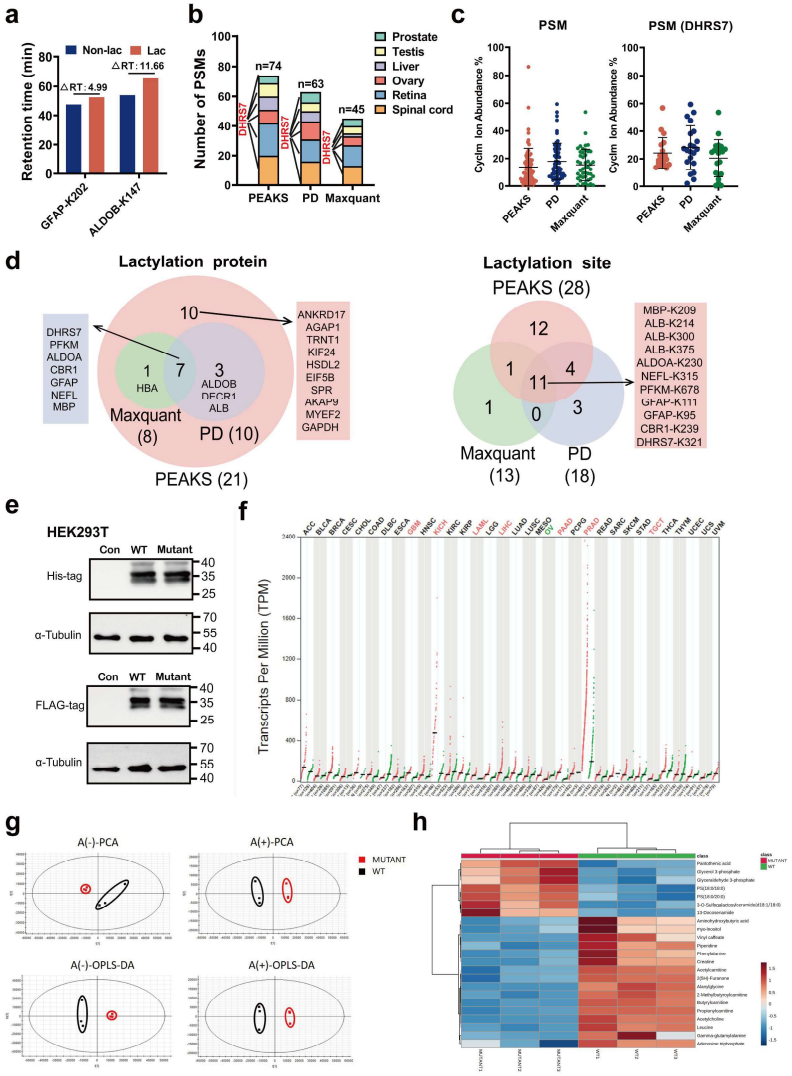

Figure S6. Benchmarking the cyclm ion to signify lactylation from the draft map of Human Proteome.

(a) Retention time shift of lactylated *vs.* non-lactylated peptide pairs identified from the Human Proteome data resources. (b) Number of lactylated PSMs identified from the proteomics data obtained from 6 tissue types retrieved by three search engines. (c) Significant ion abundance of the cyclm ion from lactylated DHRS7 peptides using data summarized in (b). (d) Venn diagram of lactylation proteins and sites identified from the 6 tissue types retrieved by three search engines. (e) Transfection efficiency of His-tagged and FLAG-tagged DHRS7 plasmid in HEK293T cells. (f) Gene expression profiles of DHRS7 across given cancer types and paired normal tissues plotted via GEPIA. (g) Principal component analysis (PCA) and orthogonal partial least square discriminant analysis (OPLS-DA) score plots of the wild type (WT) and K321A (mutant) DHRS7-overexpressing HEK293T cells using metabolomics data collected under the positive (+) and negative (-) ion modes (n=3/group). (h) Heatmap of the differentially expressed metabolites between the WT and mutant groups with the significance cutoff of $p<0.05$ (unpaired, two-tailed t-test) and VIP≥1, CV<20%.

Figure S7

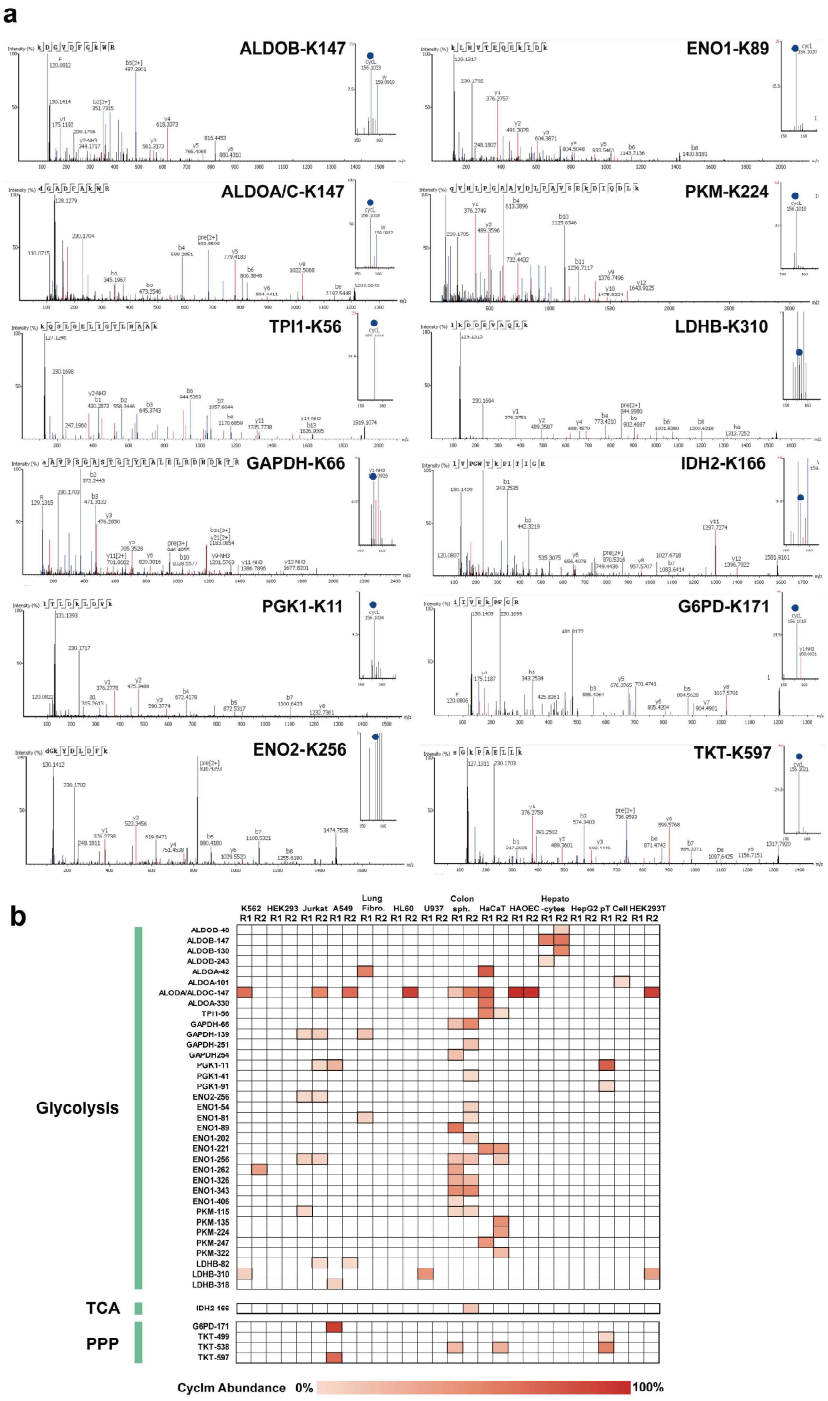

Figure S7. CycIm ion signifies enriched lactylation on glycolytic enzymes in human cells.

(a) Representative MS/MS spectra of lactylated peptides belonging to proteins involved in glycolysis including ALDOB, ALDOA/C, TPI1, GAPDH, PGK1, ENO2, ENO1, PKM2 and LDHB, proteins involved in TCA exemplified by IDH2 and proteins involved in PPP including G6PD and TKT. CycIm ion is labeled in blue in inserts. (b) Heatmap of the cyclm ion abundance in the MS/MS spectra enabling lactylation assignment for enzymes involved in glycolysis, TCA cycle and PPP across 14 cell types. Lung fibro., Lung fibroblast cells; Colon sph., Colon cancer spheroids.

Figure S8

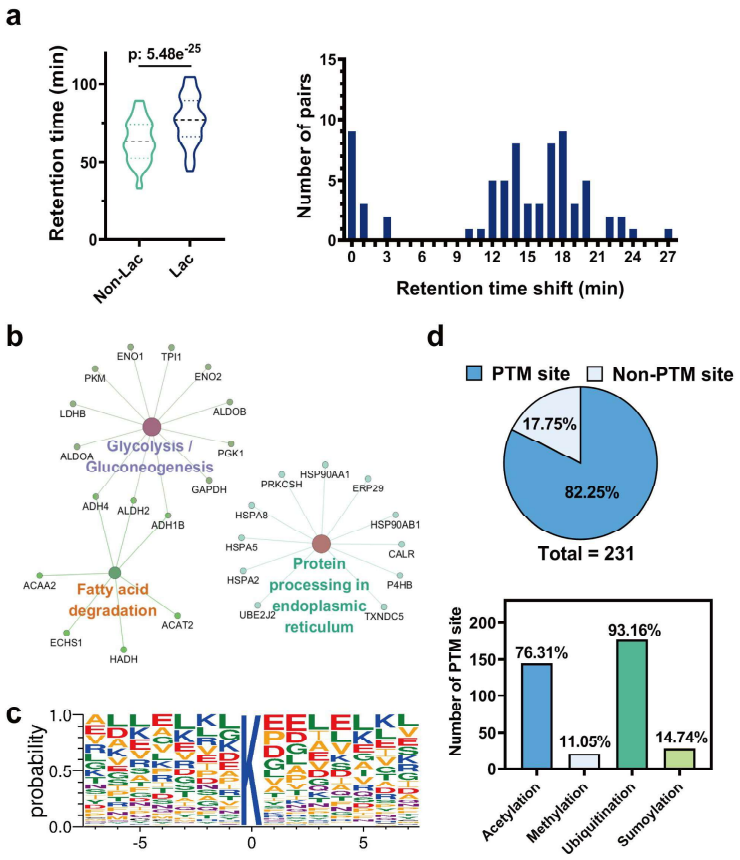

Figure S8. Functional analysis of the lactylation sites identified from the Meltome Atlas proteome data resources collected from 14 human cell types.

(a) Retention time distribution and shifts of lactylated *vs.* non-lactylated peptide pairs (n=71) identified from proteome data of 14 cell types through mining the Meltome Atlas. (b) KEGG pathway analysis of the identified lactylated proteins. (c) Motif analysis of the residues flanking the lactylated lysine residues. (d) PTM category analysis showing the distribution of previously reported PTMs on the lactylation sites uncovered in this study.

Figure S9

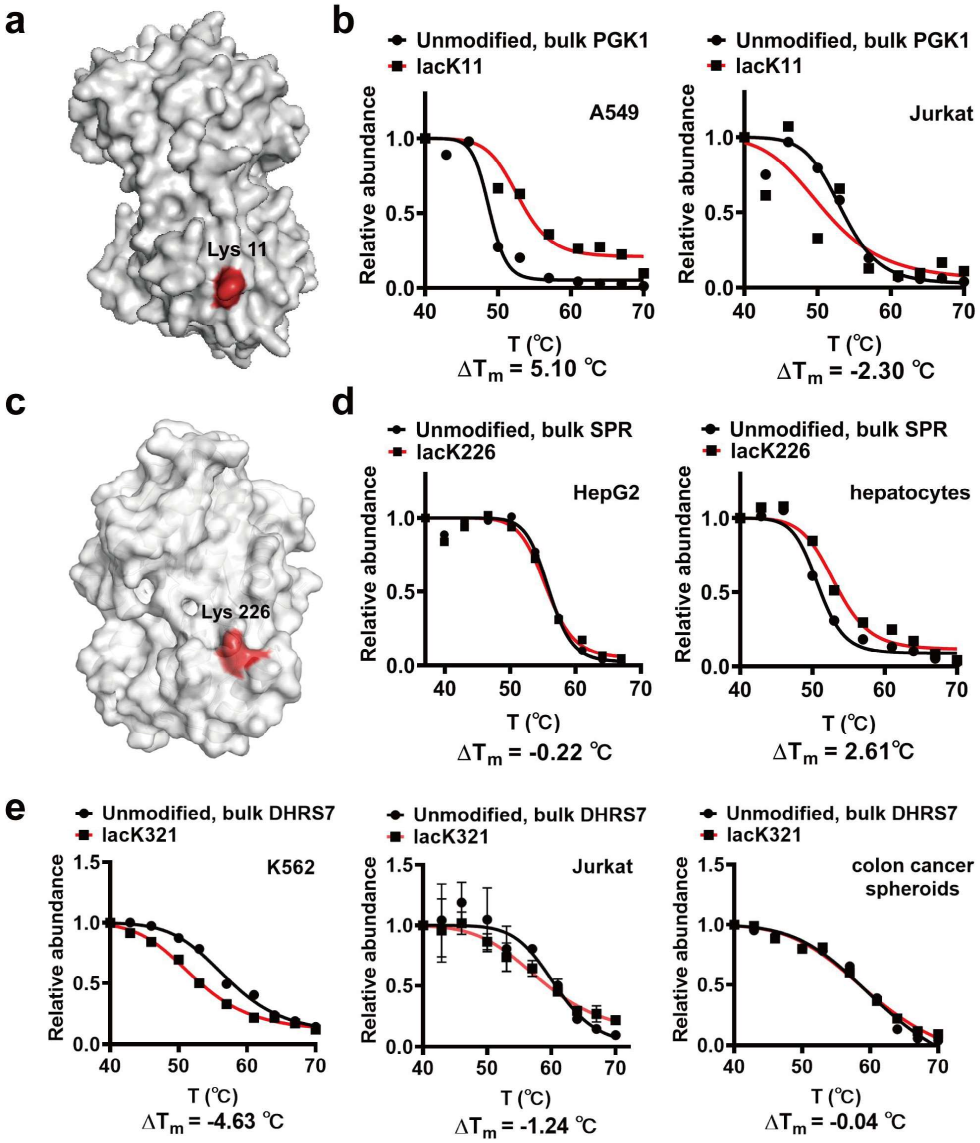

Figure S9. Melting curves of lactylated *vs.* non-lactylated human proteins reveal functional lactylation suggested by the Hotspot Thermal Profiling approach.

(a) Crystal structure of human PGK1 with highlighted lactylated lysines. (b) Melting curves of the lactylated and non-lactylated PGK1 proteoforms plotted based on the Meltome data of A549 and Jurkat cells. (c) Crystal structure of human SPR with highlighted lactylated lysines. (d) Melting curves of the lactylated and non-lactylated SPR proteoforms plotted based on the Meltome data of HepG2 cells and hepatocytes. (e) Melting curves of the lactylated and unmodified DHRS7 proteoforms based on the Meltome data of K562 cells, Jurkat cells and colon cancer spheroids.

Supporting Tables

Table 1. Summary of the lactylated peptides identified from the recombinant ENO1 protein digests using an Orbitrap and SYNAPT G2 Si Q/TOF instrument.

Table 2. Summary of the identified lactylated peptides from SILAC-processed and affinity-enriched proteome datasets collected from MCF-7 cells in response to rotenone and DCA intervention.

Table 3. Summary of the lactylated PSMs retrieved from proteome data of six human tissue types using three different database search engines.

Table 4. Summary of the lactylated PSMs identified from the 14 human cell types available in the Meltome Atlas.