

The Effects of Storage on Platelet Proteome Assessed by Protein Expression Changes and Lysine Modification Changes Using Tmt Isobaric Labeling

Li Wang

Huashan Hospital Fudan University <https://orcid.org/0000-0003-3379-2721>

Kexiang Yan

Huashan Hospital Fudan University

Zhijia Fan

Huashan Hospital Fudan University

Jie Yang

Shanghai Blood Center

Wei Liang

the Second People's Hospital of Lianyungang

Qiang Wu

Huashan Hospital Fudan University

Rufeng Xie

Shanghai Blood Center

Momoko Ishii

Yale University School of Public Health

Yuan Lu (✉ luyuan@sccl.org.cn)

Huashan Hospital Fudan University

Zhicheng Wang (✉ ahwzc@126.com)

Research

Keywords: Acetylome, Platelet, Proteome, Succinylome

Posted Date: May 15th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-28818/v1>

License: (cc) (i) This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: To provide the first insight into the proteomic dynamics during platelet storage.

Experiment design: In this study, based on TMT-labeled LC-MS/MS analysis, combined with antibody-affinity enrichment and purification for acetylated and succinylated peptides, we performed quantification of global proteomics, acetylome and succinylome. Simultaneously, dynamic molecular changes and functional transformation of platelet were also characterized under proper conditions stored for 1, 3, 5, 7 days, respectively.

Results: 3,100 proteins are quantified from a total of 3,609 proteins identified from platelets. Out of 1,308 acetylated sites identified in 648 proteins, 790 sites in 396 proteins are quantifiable. There are 1,947 succinylated sites in 959 proteins in which 1,279 sites in 661 proteins are quantifiable. We screened the differential expression changes of global proteins, acetyl- and succinyl- proteins, and systematically interpreted their molecular functions, biological processes, cellular components, pathways and motif characters to fully investigate the molecular dynamics and biological functions of platelets.

Conclusions and clinical relevance: This paper is the first systematic exploration of proteomes and modified proteomes of platelet dynamics during storage in the aim to improve our understanding of platelet biology, which may be a valuable reference for further research and clinical application.

1 Introduction

Quantitative proteomics, a powerful approach to analyze proteomic dynamics, can both identify and quantify thousands of proteins in a single sample using high-resolution liquid chromatography-tandem mass spectrometry (LC-MS/MS) [1-3]. Various isobaric mass tags are often applied for relative quantification with multiplex capabilities and high throughput [4, 5]. When primarily designed for peptide labeling, tandem mass tag has been demonstrated labeling at the protein level [6]. Tandem mass tag (TMT) [7, 8], composed of a mass reporter, a mass normalizer, and a reactive moiety, has been commercially employed for simultaneously labeling and analyzing 6, 10 or 11 samples at a time. The quantitative methods are based on amino-group labelling with isobaric TMT [9]. It has been used to quantify the relative peptide intensities of the first dimensional MS, while the second dimensional MS is to sequence peptide fragment ions for protein identification. In addition, certain enrichment and purification method, including but not limited to phosphorylation, acetylation, succinylation, and ubiquitination can be applied to modified peptides in order to achieve post-translational modification (PTM) proteome profiling [10, 11]. With various types of PTMs diversifying proteomes specifically, PTMs have been demonstrated to play important roles in regulating protein functions [12]. More widespread applications of modified “omics” have contributed to our knowledge of the complexity of biological processes mediated by PTMs [13, 14], which is critical in understanding protein dynamics and mechanisms of biological functions at the molecular level [15].

Lysine acetylation (Kac) controlled by lysine acetyltransferases is a reversible process of PTM in biological organisms[16]. Recent advances to identify and characterize Kac have improved our understanding of its biological significance[17-19]. To date, acetylomic analyses based on LC-MS/MS have unveiled roles of Kac events associated with all cellular processes, ranging from gene expression and cell signaling to metabolism[20, 21]. However, Kac dynamics have not been well studied during platelet storage to identify all Kac sites and describe their biological functions.

Proteins are frequently modified by acylations other than Kac, such as succinylation, one of an important lysine modification process. Lysine succinylation (Ksucc) was first identified in *Escherichia coli*[22]. Although it is a novel PTM, Ksucc is evolutionarily conserved in multiple species[23-25], involved in the regulation of a number of essential cellular functions including metabolic processes, transcription, translation and others [20]. Interestingly, protein lysine succinylation is preferred in protein biosynthesis and carbon metabolisms, such as the tricarboxylic acid (TCA) cycle and fatty acid metabolism[26]. Even though Ksucc has been broadly investigated and validated in various organisms[27], limited has been explored in the context of platelet storage, which may involve substantial metabolic changes.

Platelet quality is an essential index for transfusion in the clinic[28, 29]. The activity of platelets is rapidly declining in fresh blood, especially at 4°C, leading to irreversible changes of the platelet membranes such that only 40% of overall activities are preserved after 6 hours, and 20% after 12 hours. At less than ambient temperature, there are significant changes in platelets[30], including phenotypical change from stationary disk to poly-pseudopod, increased filamentous actin, depolymerization of platelet microtubules, sharply higher calcium ions, and secretion and fusion of lysosomes and so on[31, 32]. On the other hand, warmer temperatures may increase bacterial risk, making platelets susceptible to infections[33]. All aforementioned changes can dramatically reduce physiological activity, leading to ineffective platelet therapy, intoxication or death[34]. So, it is pivotal to preserve platelets using proper techniques[35, 36]. Notably, it is also essential to analyze molecular changes during platelet storage, albeit it has not yet been accomplished so far.

For the first time, stored platelets (22 ± 2°C, with agitation on nutator) are quantified by integrating global proteome, acetylome and succinylome for 1, 3, 5, 7 days. We identified 3,609 proteins and quantified 3,100 of them: 1,308 Kac sites matched on 648 Kac proteins and quantifiable 790 Kac sites identified on 396 Kac proteins; 1,947 Ksucc sites matched on 959 Ksucc proteins and quantifiable 1,279 Ksucc sites identified on 661 Ksucc proteins. By comparing differential protein expressions with an average cut-off change of 1.3-fold, we systematically analyzed molecular changes and functional transformations (p-value ≤ 0.05). Bioinformatic analysis revealed that these differential proteins are involved in important cellular processes and distributed in diverse subcellular compartments. This study provides the first insight into the proteomic dynamics of platelet storage, as well as biological relevance of Kac and Ksucc events, increasing our understanding of platelet biology and transfusion safety in clinical settings.

2 Materials And Methods

2.1 Clinical samples of the platelets.

Platelet samples were taken from Shanghai Blood Center and stored under standard clinical conditions[37]. Platelet samples were used at the 1st, 3rd, 5th and 7th day of storage for procedures described below. For studies involving human subjects, approval was obtained from the Huashan Hospital Institutional Review Board (HIRB), Fudan University, Shanghai, China. Informed consent was provided in accordance with the Declaration of Helsinki.

2.2 Protein extraction from platelet and trypsin digestion.

Sample debris was removed by 20,000g centrifugation at 4°C for 10 minutes and the subsequent protein concentration was tested using the BCA Quant KIT (GE Healthcare Life Sciences™, BJ, CN.) according to the manufacturer's instructions. For digestion, the protein solution was reduced with 10mM DL-dithiothreitol (DTT) (Sigma-Aldrich, SH, CN.) for 1 hour at 37 °C and then alkylated with 20mM 3-iodoacetamide (IAA) (Sigma-Aldrich, SH, CN.) for 45 minutes at room temperature in dark. For trypsin digestion, the protein sample was diluted by adding 100mM tetraethylammonium tetrahydroborate (TEAB) to urea concentration less than 2M. Finally, trypsin (Promega, Madison, WI, USA.) was added at 1:50 trypsin-to-protein mass ratio for the first digestion overnight and 1:100 trypsin-to-protein mass ratio for a second 4-hour digestion. Approximately 100µg protein for each sample was digested with trypsin (Promega, Madison, WI, USA.) for the following experiments (Kac and Ksuss included).

2.3 TMT labeling and HPLC fractionation.

After trypsin digestion, the peptide was desalted by Strata X C18 SPE column (Phenomenex, TJ, CN.) and vacuum-dried. The peptide was reconstituted in 0.5M TEAB and processed according to the manufacturer's protocol for TMTsixplex™ Isobaric Label Reagent Set (ThermoFischer Scientific, SH, CN.). In brief, one unit of TMT reagent (defined as the amount of reagent required to label 100µg of protein) was used and reconstituted in 24µl acetonitrile (ACN) (Fisher Chemical, SH, CN.). The peptide mixtures were then incubated for 2 hours at room temperature and pooled, desalted and dried by vacuum centrifugation.

The sample was then fractionated by high pH reverse-phase HPLC using Agilent 300 Extend C18 column (5µm particles, 4.6mm ID, 250mm length). Briefly, peptides were first separated with a gradient of 2% to 60% ACN in 10mM ammonium bicarbonate pH 10 over 80min into 80 fractions. Then, the peptides were combined into 18 fractions and dried by vacuum centrifugation.

2.4 Affinity enrichment for Kac and Ksu peptides.

For Kac and Ksucc peptide enrichment, tryptic peptides were dissolved in NETN buffer (100mM NaCl, 1mM EDTA, 50mM Tris-HCl, 0.5% NP-40, pH 8.0), and then incubated with pre-washed antibody beads (PTM Biolabs, Chicago, IL, USA.) at 4°C overnight with gentle shaking. The beads were washed four times with NETN buffer and twice with ddH₂O (ThermoFischer Scientific, SH, CN.). The bound peptides were eluted from the beads with 0.1% trifluoroacetic acid (TFA) (Sigma-Aldrich, SH, CN.). The eluted fractions were combined and vacuum-dried. The resulting peptides were cleaned with C18 ZipTip[®] pipette tips (EMD Millipore, SH, CN.) according to the manufacturer's instructions, followed by LC-MS/MS analysis.

2.5 LC-MS/MS analysis for global proteome, acetylome, and succinylome of the platelets.

Peptides were dissolved in 0.1% FA, directly loaded onto a reversed-phase pre-column, AcclaimTMPepMapTM 100 C18(ThermoFischer Scientific, SH, CN.). Peptide separation was performed using a reversed-phase analytical column, AcclaimTMPepMapTM RSLC (ThermoFischer Scientific, SH, CN.). The gradient comprised of an increase of solvent B (0.1% formic acid in 98% ACN) from 8% to 25% for the first 20 minutes, 25% to 40% for the subsequent 12 minutes, then increasing to 80% in the next 4 minutes, and maintaining at 80% for the last 4 minutes. All were performed at a constant flow rate of 400 nl/min on an EASY-nLC 1000 UPLC system, the resulting peptides were analyzed by QExactiveTMPlus Hybrid Quadrupole-OrbitrapTM Mass Spectrometer (ThermoFisher Scientific, SH, CN.).

The peptides were subjected to NanoSpray Ionization (NSI) Source followed by tandem mass spectrometry (MS/MS) in Q ExactiveTMPlus (ThermoFischer Scientific, SH, CN.) connected online to the UPLC. Intact peptides were detected in the Orbitrap at a resolution of 70,000. Peptides were selected for MS/MS using normalized collision energy (NCE) setting as 30; ion fragments were detected in the Orbitrap at a resolution of 17,500. A data-dependent procedure that alternated between one MS scan followed by 20 MS/MS scans was applied for the top 20 precursor ions above a threshold ion count of 5E3 in the MS survey scan with 15.0s dynamic exclusion. The electrospray voltage applied was 2.0 kV. Automatic gain control (AGC) was used to prevent overfilling of the Orbitrap; 5E4 ions were accumulated for the generation of MS/MS spectra. For MS scans, the m/z scan range was 350 to 1800. The fixed first mass was set as 100 m/z.

2.6 Database search and QC validation of MS data.

The resulting MS/MS data was processed using MaxQuant[38] with an integrated Andromeda search engine (v.1.5.2.8). Tandem mass spectra were searched against the SwissProt human database concatenated with reverse decoy database. Cleavage agent Trypsin/P allowed up to 4 missing cleavages, 5 modifications per peptide and 5 charges. A mass error was set to 10 ppm for precursor ions and 0.02 Da for the fragment ions. Cysteine Carbamidomethylation was specified as fixed modification while oxidation on Methionine, acetylation on Lysine, and acetylation on protein N-terminal were specified as

variable modifications. False discovery rate thresholds for protein, peptide and modification site were specified at 1%. Minimum peptide length was set at 7. TMT-6-plex was selected as the quantification method of proteins by report ion in MS/MS. All the other parameters in MaxQuant were set to default values. The site localization probability was set as > 0.75 . For global proteomics, acetylation and Succinylations, Cysteine Carbamidomethylation was specified as fixed modification while oxidation on Methionine, and acetylation or succinylation on Lysine, acetylation on protein N-terminal, respectively.

Validation of the MS data was done using mass error distribution of all identified peptides and peptide length distribution. Firstly, we checked the mass error of all the identified peptides. The distribution of mass error is near zero and most of them were less than 0.02 Da which indicates that the mass accuracy of the MS data meets the requirement. Secondly, the length of most peptides was distributed between 8 and 20, which agreed with the property of tryptic peptides, indicating that sample preparation has reached the quality standard.

2.7 Bioinformatics annotation of differential proteins.

Gene Ontology (GO) annotation[39] proteome was derived from the UniProt-GOA database (www.ebi.ac.uk/GOA/). Firstly, protein ID was converted into UniProt ID and the converted UniProt ID was mapped into GO IDs based on protein ID. If some identified proteins were not annotated by UniProt-GOA database, the InterProScan software would be used to annotate protein's GO functions based on protein sequence alignment method. Then proteins were classified by Gene Ontology annotation according to three categories: biological process, cellular component and molecular function. For each category, a two-tailed Fisher's exact test was employed to test the enrichment of the differentially expressed protein against all identified proteins. The GO with a corrected p-value < 0.05 is considered significant.

Kyoto Encyclopedia of Genes and Genomes (KEGG) database[40, 41] was used to annotate protein pathway. Firstly, KEGG online service tools, KAAS was used to annotate protein's KEGG database description. Then KEGG Mapper, a KEGG online service tool, was used to map the aforementioned annotation result. A two-tailed Fisher's exact test was applied to test the enrichment of the differentially expressed protein against all identified proteins. The pathway with a corrected p-value < 0.05 was considered significant. These pathways were classified into hierarchical categories according to the KEGG website.

2.8 Functional enrichment of differential proteins.

KEGG database was used to identify enriched pathways by a two-tailed Fisher's exact test to test the enrichment of the differentially expressed protein against all identified proteins. Correction for multiple hypothesis testing was carried out using False Discovery Rate (FDR) control methods[42]. The pathway with a corrected p-value < 0.05 was considered significant.

2.9 Motif analysis and enrichment-based clustering of protein express profiling.

Soft motif-x[43] was applied to calculate the model of amino acid sequences in specific positions of modify-21-mers (10 amino acids upstream and downstream of the site) in all protein sequences. All the database protein sequences from Sections 2.7 and 2.8 were used as a background database parameter and other parameters were set to default.

All the substrate categories obtained after enrichment were collated along with their p-values, and then filtered for those categories which were at least enriched in one of the clusters with a lower p-value threshold of 0.03. This filtered p-value matrix was transformed by the function $x = -\log_{10}(\text{p-value})$. At last, these x values were z-transformed for each category. These z scores were then clustered by one-way hierarchical clustering (Euclidean distance, average linkage clustering) in Genesis. Cluster membership was visualized by a heat map using the “heatmap.2” function from the “gplots” R-package.

3 Results

3.1 Generation of the global proteome, lysine acetylome and succinylome for platelet.

To comprehensively profile the multi-omics of platelet and illustrate the dynamics during platelet storage, we simultaneously identified and quantified the global proteome, acetylome, and succinylome profiles at the 1st, 3rd, 5th and 7th days for platelets stored under standard conditions in three replicates. We applied TMT-labeled LC-MS/MS with HPLC prefractionation, combined with antibody-based immuno-affinity purification to enrich Kac and Ksucc peptides. The overall strategy was illustrated in Fig. 1.

We then achieved comprehensive profiling that includes both identification and quantification, as shown in Table 1. For the global proteome, 3,609 proteins were identified with 3,100 quantitative proteins in total (See Supplemental Table 1 for global proteome). For acetyl-proteome after acetylated peptide enrichment, 1,308 Kac sites were matched on 648 identified proteins, among which 790 sites in 396 proteins were quantified (See Supplemental Table 2 for acetylome). For succinylome altogether, 1,947 Ksucc sites in 959 proteins were identified, with 1,279 quantitative sites in 661 proteins (See Supplemental Table 3 for succinylome). From the comprehensive profiling of proteins in platelets, the quantities of acetyl-proteins and succinyl-proteins were 18.0% and 26.6%, respectively. This indicates that succinylated modification may be more extensive than acetylation in platelet.

Table 1
Comprehensive profiling of global proteome,
acetylome, and succinylome.

Profiling Names	Identification	Quantification
Global Proteins	3,609	3,100
Kac Sites	1,308	790
Kac Proteins	648	396
Ksucc Sites	1,947	1,279
Ksucc Proteins	959	661

Table 2
Differentially quantified proteins in global proteome, acetylome, and succinylome. Note: $p < 0.05$.

Group Comparison	Type	Global Proteins		Kac Proteins		Ksucc Proteins	
		Up-regulated (>1.3)	Down-regulated (<0.77)	Up-regulated (>1.3)	Down-regulated (<0.77)	Up-regulated (>1.3)	Down-regulated (<0.77)
D3/D1	Sites			14	114	0	735
	Proteins	90	39	13	79	0	433
D5/D1	Sites			4	100	21	9
	Proteins	86	30	4	74	20	9
D5/D3	Sites			28	50	778	0
	Proteins	23	39	26	34	453	0
D7/D1	Sites			96	70	75	11
	Proteins	27	29	67	53	62	10
D7/D3	Sites			214	0	854	0
	Proteins	18	74	140	0	478	0
D7/D5	Sites			192	0	58	12
	Proteins	58	93	125	0	53	12

Table 3
The feature sequence near acetylated sites and the statistics by motif analysis.

Motif	Motif Score	Foreground		Background		Fold Increase
		Matches	Size	Matches	Size	
.....KK.....	16	241	1261	47857	602429	2.41
.....KR.....	16	136	1020	33799	554572	2.19
.....K.K.....	16	148	884	40539	520773	2.15
.....KH.....	16	84	736	13493	480234	4.06
.....KN.....	13.52	81	652	22772	466741	2.55
.....K K.....	10.15	90	571	33951	443969	2.06
.....K.R.....	8.72	62	481	23183	410018	2.28
.....KF.....	8.53	34	234	13624	298960	3.19
.....K.E.....	8.37	73	419	32900	386835	2.05
.....KS.....	7.81	63	346	30612	353935	2.11
.....KT.....	7.34	49	283	24363	323323	2.3

To validate the quality of MS and MS/MS profiling, we evaluated the mass error of the all identified peptides. The distribution of mass error was near zero and most of them were less than 0.02 Da, demonstrating that the overall accuracy of the MS data meets the requirement (Supplemental Fig. 1A). The length of most peptides was distributed between 8 and 20, which agrees with the property of tryptic peptides (Supplemental Fig. 1B), also indicating that the sample preparation reached the required standard. To assess the reproducibility among the 3 biological replicates, we performed the repeatability analysis by Pearson correlation coefficient. As shown in Supplemental Fig. 2, regardless of profile types (global proteome, acetylome or succinylome), the Pearson correlation coefficients of biological samples in triplicates were all above 0.8, while the correlations between different groups were poor, indicating good biological repeatability within every group of the experiments. 6 typical MS/MS spectra images of 3 Kac and 3 Ksucc peptides were also showed in Supplemental Fig. 2D. Together, the proteomic analysis was robust in which all data can be used for the subsequent analyses with high quality.

3.2 Comparison of differential protein expression of stored platelet for 1, 3, 5, 7 days.

By comparing the differential protein expression with an average cut-off change of 1.3-fold and p-value 0.05, we screened the significantly differential proteins among the groups of Day 1 (D1), Day 3 (D3), Day 5 (D5), and Day 7 (D7). As shown in Table 2 and Fig. 2 for the global proteins, 90 proteins were up-

regulated and 39 proteins were down-regulated in group D3/D1; 86 proteins were up-regulated and 30 proteins were down-regulated in group D5/D1, and 23 up-regulated proteins and 39 down-regulated proteins in group D5/D3; 27 proteins were up-regulated and 29 proteins were down-regulated in group D7/D1, and 18 up-regulated proteins and 74 down-regulated proteins in group D7/D3, and 58 up-regulated proteins and 93 down-regulated proteins in group D7/D5(See Supplemental Table 4 for differential proteins).

Table 4
The feature sequence near succinylated sites and the statistics by motif analysis.

Motif	Motif	Foreground		Background		Fold Increase
	Score	Matches	Size	Matches	Size	
.....E KR.....	32	60	1245	3448	546682	7.64
.....E KK.....	32	101	1879	5366	602429	6.07
.....D KK.....	29.79	49	1694	2500	591910	6.85
.....RKY.....	28.49	27	1298	921	548954	12.4
.....KKY.....	27.37	28	1185	1335	543234	9.61
.....K.KE.....	26.44	40	492	4268	399669	7.61
.....KHE.....	25.76	26	1778	816	597093	10.7
.....DKY.....	25.36	20	822	754	481039	15.52
.....K.K.D.....	25.1	44	1342	2198	551152	8.22
.....DKK.....	24.99	28	1370	1520	552672	7.43
.....R.K.E.....	24.42	40	1410	2699	555371	5.84
.....KDD.....	23.92	26	1271	1351	548033	8.3
.....RKH.....	23.52	20	1128	1015	539686	9.43
.....K.K.E.....	23.14	58	1752	4367	596277	4.52
.....LKK.....	22.86	40	1108	3999	538671	4.86
.....K.KD.....	22.5	29	1157	2213	541899	6.14
.....KKH.....	22.16	16	508	1137	400806	11.1
.....KK.....	16	94	452	27765	395401	2.96
.....K.D.....	16	103	925	20417	501456	2.73
.....KR.....	16	94	602	24722	425528	2.69
.....KP.....	16	235	1645	34039	589410	2.47
.....K.E.....	16	143	1068	33216	534672	2.16
.....RK.....	14.3	101	703	28013	453641	2.33
.....KD.....	12.63	99	802	26744	480285	2.22

Among the Kac proteins, 14 Kac sites in 13 proteins were up-regulated and 114 Kac sites in 79 proteins were down-regulated in group D3/D1; 4 Kac sites in 4 proteins were up-regulated and 100 Kac sites in 74

proteins were down-regulated in group D5/D1; 28 up-regulated Kac sites in 26 proteins and 50 down-regulated Kac sites in 34 proteins in group D5/D3; 96 Kac sites in 67 proteins were up-regulated and 70 Kac sites in 53 proteins were down-regulated in group D7/D1; 214 up-regulated Kac sites were present in 140 proteins with no down-regulated Kac sites in group D7/D3; finally, 192 up-regulated Kac sites in 125 proteins and no down-regulated Kac sites in group D7/D5 (See Supplemental Table 5 for differential Kac proteins).

For the differential Ksucc proteins, no Ksucc sites were up-regulated and 735 Ksucc sites in 433 proteins were down-regulated in group D3/D1; 21 Ksucc sites in 20 proteins were up-regulated and 9 Ksucc sites in 9 proteins were down-regulated in group D5/D1; 778 up-regulated Ksucc sites in 453 proteins and no down-regulated Ksucc sites in group D5/D3; 75 Ksucc sites in 62 proteins were up-regulated and 11 Ksucc sites in 10 proteins were down-regulated in group D7/D1; 854 up-regulated Ksucc sites in 478 proteins and no down-regulated Ksucc sites in group D7/D3; finally, 58 up-regulated Ksucc sites in 53 proteins and 12 down-regulated Ksucc sites in 12 proteins in group D7/D5 (See Supplemental Table 6 for differential Ksucc proteins).

All the differences listed above are direct reflection of functional changes that gradually occurred during platelet storage. Interestingly, when D3 is compared to D1, up-regulated proteins were more than down-regulated ones, while it was opposite for the Kac and Ksucc proteins comparison. On the contrary, even though the total down-regulated proteins were a little more than the up-regulated ones, the up-regulated Kac and Ksucc proteins were much more than the corresponding down-regulated ones in the group of D7/D1, D7/D3 and D7/D5, especially D7/D3 and D7/D5. The shift of up- and down-regulated proteins and their modifications may be related to their functional transformation during platelet storage. The biological significance of this shift was carefully carried out to annotate the differential proteins as follows.

3.3 Dynamic characterization of platelet storage by functional annotation and enrichment analysis.

We tried to comprehensively annotate differential expressed proteins based on several categories, such as Gene Ontology (GO), KEGG pathway and subcellular localization. We found a wide range of functional distribution of the up- or down-regulated proteins in the comparison groups. To further dive in the difference among the desired storage time points, we screened and selected significant entries by performing functional enrichment of differential qualified proteins (Fig. 3). The most significantly enriched biological processes of up-regulated proteins were humoral immune response, adaptive immune response and complement activation in D3/D1, D5/D1 and D7/D1, which might suggest immunity was activated during storage. Other molecular characteristics, such as rapid increase in peptidase activity, antigen binding, and transporter activity, have also been noted. On the contrary, many down-regulated proteins were enriched in the regulation of coagulation and hemostasis, response to bacterium and peptide biosynthetic process. It might indicate that platelet activity declined over storage time.

Surprisingly, when comparing D7/D3 and D7/D5, there was a sudden decrease in cell immunity and complement activation, but also an increase in cellular detoxification and xenobiotic catabolic process. This might indicate a slight drop of immune response at the end of storage, when a large number of toxins had been accumulated. For differential Kacproteins, significant increase of metabolic and catabolic processes was observed, which involved small molecules, nucleic acid, and various compounds. The most enriched cellular components were mitochondria and cell junctions (See Supplemental Table 7). The enrichment of up-regulated Ksuccproteins mainly focused on cell junction assemblies, cytoplasmic vesicles and metabolic processes, which was similar to the up-regulated Kacproteins. Whereas the enrichment of down-regulated proteins was reflected in protein localization, secretion and the cytoplasmic region (See Supplemental Table 8).

Analysis in KEGG pathway-based enrichment showed significant increases of complement and coagulation cascades, *Staphylococcus aureus* infection, systemic lupus erythematosus, and pertussis during the first 5 days of storage, while ribosome and cytokine-cytokine receptor interaction dropped (Fig. 4). At the end of storage, almost all aforementioned observations decreased. However, we also observed elevations of drug and xenobiotics metabolisms induced by cytochrome P450, chemical carcinogenesis, and oxidative phosphorylation. It is also important to notice the large variations in differential Kacproteins expressions. The up-regulated Kac proteins were enriched in metabolic pathways, focal adhesion and TCA cycle. On the other hand, down-regulated Kac proteins focused on platelet activation and fatty acid metabolism, which might signify a decline in the stored platelet activity (See Supplemental Fig. 3). The down-regulated Ksucc proteins mainly enriched in glycolysis/gluconeogenesis and focal adhesion on D3/D1 and D5/D1; At the end of storage, focal adhesion and regulation of actin cytoskeleton were clearly enriched for the up-regulated proteins (See Supplemental Fig. 4).

Analysis above might suggest that platelet immune responses were activated and processing during the storage, as well as complement and coagulation cascades. Cellular detoxification and cell junction also increased over time. On the contrary, it seemed like that platelet membranes and ribosomes were slightly suffering from damage. The enrichment of differential Kac and Ksucc proteins was in line with the global proteins. Altogether, the enrichment analysis of global proteins, Kac and Ksucc proteins gave us a clue for the dynamic changes of the stored platelet biology.

3.4 Dynamic clustering of differential protein expression.

Based on the clustering analysis, the differential proteins were divided into 12 clusters (Fig. 5). Every cluster had specific traits among the groups, indicating differential patterns for the protein expressions of platelets among the 4 groups of storage time. There were 9 clusters for the differential Kac proteins and 5 clusters of Ksucc proteins (Supplemental Fig. 5 and Supplemental Fig. 6).

According to the dynamic clustering, we obtained heatmaps from GO and KEGG pathway enrichment analysis (Fig. 6). The enriched biological processes for global proteins varied in differential groups, same as the KEGG pathway, suggesting dynamic protein expressions and regulation patterns during platelet

storage. For example, regulation of lipid metabolic process clearly increased from D1 to D3, and continued increase was observed until D5, but declined on the 7th day, indicating that the metabolic activity enhanced at the start due to sudden stimulation of microenvironmental changes, and then decreasing to a complete loss of activity *ex vivo*. The same regularity was applicable to other metabolic processes or transport of metabolites. Some pathways may be initiated sooner upon collection of platelets, and then decline gradually with complement and coagulation cascades; however, other pathways may be activated later, such as drug metabolism induced by cytochrome P450 and chemical carcinogenesis among others. The heatmap patterns of Kac protein enrichment followed the same regularity as that of global proteins, and minor differences lied in the specific biological processes and pathways involved in Kac regulation (See Supplemental Fig. 7). Interestingly, Ksucc protein enrichment heatmaps also contained a different set of biological processes and pathways as those of Kac protein enrichment (See Supplemental Fig. 8). Upon comparison of heatmaps, Kac and Ksucc proteins were noted to be involved in regulations of different biological processes and pathways with minor overlaps, indicating the great differences of two modifications. Importantly, they also complemented with the global protein expression patterns, meaning that Kac and Ksucc were indispensable for the regulations of protein functions.

3.5 Kac and Ksucc motifs identified from Kac and Ksucc sites of stored platelets.

To reveal potential regularity of amino acid sequences in front of and behind Kac or Ksucc sites, motif analysis was utilized to calculate probabilities of amino acids near a specific Kac or Ksucc site. Amino acids at the foreground and background of Kac or Ksucc sites from the whole site were matched and converted into motif scores and fold changes, in order to speculate potential amino acids near the Kac or Ksucc sites and obtain statistical significance. As shown in Table 3, most likely 11 sequences near the Kac site were obtained, where KK, KR, KK, and KH had the highest motif scores. Similarly, most likely 23 sequences near the Ksucc site were summarized in Table 4: EKR, EKK, DKK, and RKY were the most possible motifs around the Ksucc site.

To illustrate quantitative differences of the sequence features near specific Kac or Ksucc sites, we displayed motif enrichment for upstream and downstream amino acids of Kac and Ksucc sites as heatmaps (See Fig. 7). The most possible amino acid after Kac site was one of H, K or R, and the second most likely was K; while the first most possible amino acid before Kac site was D or N, and the second most likely was A, and so forth. For the Ksucc site, the most possible amino acid after the site was one of D, H, K, P, R or Y, and the second most likely was D or E; while the most possible amino acid before the site was D, K, R or Y, and the subsequently likely one was K, and so on.

4 Discussion

In the “-omics” era, it is desirable to identify and quantify the entire suite of expressed proteins and uncover the dynamic changes they undergo during a process of interest[44, 45]. The next-generation proteomics has been greatly improved to achieve proteomic dynamics in depth, making it possible to systematically describe molecular changes in a cell and those of tissue proteomes with a high temporal resolution[46, 47]. Recent studies have reported that platelet functions, such as granule release, adhesion and aggregation could be affected by acetylation. Those regulation has been demonstrated in the presence of p300, HDAC6, and Sirtuins in platelets. Knowing the acetylation mechanisms in platelets is related to the treatment and prevention of cardiovascular diseases and will open new possibilities for regulating platelet functions [48].

The TMT-labeled LC-MS/MS combined with pre-fractionation by HPLC allows for comprehensive proteome profiling with deep coverage[49, 50]. Combined with acetylated and succinylated peptide enrichment methods[51-53], this strategy was able to determine the level of global proteomic changes, as well as expression changes of acetylome and succinylome, simultaneously[54]. To analyze the dynamics and regulations of platelets stored for 1, 3, 5, 7 days, we applied the strategy described above to achieve global proteome, acetylome, and succinylome for clinical platelets at 4 time points, respectively. In total, 3,609 proteins were identified, and among them, 3,100 proteins were quantified. 1,308 Kac sites in 648 proteins were identified with quantitative 790 sites in 396 proteins, as well as 1,947 Ksucc sites in 959 proteins with quantitative 1,279 sites in 661 proteins.

Based on our high-resolution data, a systematic analysis of protein expressions and modifications upon acetylation and succinylation were followed to illustrate dynamic changes of platelet biology during the storage time. Firstly, we screened the differential proteins, acetyl-proteins and succinyl-proteins between the two groups, separately. Next, based on the differentiation of protein expressions and regulations, respective biological significance was revealed to identify and understand the processes in which the stored platelets were undergoing. Finally, we also characterized the motifs close in proximities to the acetyl- and succinyl-sites, indicating the sequence characters near specific Kac or Ksucc sites to predict their targets and functions.

For the global protein expression, the most significant changes were the activation of immune responses as adaptation to the microenvironment *ex vivo*, including humoral immune response, adaptive immune response, and lymphocyte-mediated immunity, all of which were highly consistent with observations reported by previous research[12]. Complement activation and protein activation cascades were also evidently stimulated. On the contrary, regulations of hemostasis, cell division, and biosynthetic processes were reduced. For example, the growth factor adapter protein (p66Shc), a 66-kDa isoform of ShcA, was reported a role in ROS generation, which resulting in the oxidative stress during platelet storage. Our proteomic data show that p66Shc expressions increased on the later phase of platelet storage, which gave us inspiration to probe the mechanisms under platelet storage lesions in our recently published paper[55].

Lysine is a common target for post-translational modification, and acetylation is the most extensive pattern involved in efficient biological mechanism for protein regulation[20, 56]. During the process of platelet storage, the Kac proteins were participating in a wide range of biological functions and regulations, such as involvement in metabolic pathways, focal adhesion, cell junction, and platelet activation among others. Beyond acetylation, lysine succinylation has also been extensively explored, involved in glycolysis/gluconeogenesis, protein processing in the endoplasmic reticulum, regulation of cell motility, as well as platelet activation. The biological changes, as well as protein regulations, manifested the transformation that platelets were undergoing to shape the dynamics of platelet biology.

5 Conclusion

In summary, this is the first systematic exploration of proteome and modification proteome for platelet dynamics during storage. With comprehensive proteome, acetyl- and succinyl-proteome analyses, we screened changes in the expressions of global proteins, Kac and Ksucc proteins, and systematically analyzed the molecular dynamics and biological functions. By interpreting differential changes of molecular functions, biological processes, cellular components, and pathways, we observed that immune responses were stimulated for the first 5 days followed by decline and loss of activity. This was also similar to the observations of complement and coagulation cascades, some metabolic pathways for biosynthesis and degradation during the storage of platelet. In particular, ribosomal and cellular membranes of platelets were continuously and irreversibly damaged once out of the body, while chemical carcinogenesis was gradually accumulated with time *ex vivo*, indicating accumulation of hazardous substances with increasing platelet storage time. The systematic differentiation analyses of global proteome and acetyl-proteome improved our molecular understanding of platelet biology, which may be a valuable reference for biological and clinical application. In addition, these findings both elucidated platelet dynamics and complexity of protein regulations. Further research is needed to investigate biological significance of protein modifications.

Declarations

Ethical Approval and Consent to participate

For studies involving human subjects, approval was obtained from the Huashan Hospital Institutional Review Board (HIRB), Fudan University, Shanghai, China. Informed consent was provided in accordance with the Declaration of Helsinki.

Availability of data and materials

Available

Competing interests

All authors have read and approved the final manuscript. All authors declare no financial/commercial conflicts of interest.

Funding

This study was supported by grants from the National Natural Science Foundation of China (NSFC 81673080, 81571365, and 81270650).

Authors' contributions

Zhicheng Wang and Yuan Lu contributed to the design of this investigation. The experiments were performed by Li Wang, Zhijia Fan, Kexiang Yan, and Rufeng Xie. Analysis and interpretation of results were completed by Zhicheng Wang, Yuan Lu, Jie Yang, Wei Liang, and Qiang Wu. This manuscript was written by Li Wang and Zhicheng Wang.

Acknowledgements

We thank Xiang Gao and Yankuan Zhou (Jingjie PTM Biolabs (Hangzhou) Co., Ltd., Hangzhou, Zhejiang, China) for technical support with LC-MS/MS and bioinformatics analysis of platelet samples.

Authors' information

See on the first page of this manuscript.

References

1. Gallien S, Domon B: **Advances in high-resolution quantitative proteomics: implications for clinical applications.***Expert Rev Proteomics* 2015, **12**:489-498.
2. Li H, Han J, Pan J, Liu T, Parker CE, Borchers CH: **Current trends in quantitative proteomics - an update.***J Mass Spectrom* 2017, **52**:319-341.
3. Chen Y, Wang F, Xu F, Yang T: **Mass Spectrometry-Based Protein Quantification.***Adv Exp Med Biol* 2016, **919**:255-279.
4. Chahrour O, Cobice D, Malone J: **Stable isotope labelling methods in mass spectrometry-based quantitative proteomics.***J Pharm Biomed Anal* 2015, **113**:2-20.
5. Liang HC, Lahert E, Pike I, Ward M: **Quantitation of protein post-translational modifications using isobaric tandem mass tags.***Bioanalysis* 2015, **7**:383-400.
6. Sinclair J, Timms JF: **Quantitative profiling of serum samples using TMT protein labelling, fractionation and LC-MS/MS.***Methods* 2011, **54**:361-369.

7. Hung CW, Tholey A: **Tandem mass tag protein labeling for top-down identification and quantification.***Anal Chem* 2012, **84**:161-170.
8. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C: **Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS.***Anal Chem* 2003, **75**:1895-1904.
9. Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez JC: **Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags.***Anal Chem* 2008, **80**:2921-2931.
10. Ke M, Shen H, Wang L, Luo S, Lin L, Yang J, Tian R: **Identification, Quantification, and Site Localization of Protein Posttranslational Modifications via Mass Spectrometry-Based Proteomics.***Adv Exp Med Biol* 2016, **919**:345-382.
11. Rogers LD, Overall CM: **Proteolytic post-translational modification of proteins: proteomic tools and methodology.***Mol Cell Proteomics* 2013, **12**:3532-3542.
12. Liu J, Qian C, Cao X: **Post-Translational Modification Control of Innate Immunity.***Immunity* 2016, **45**:15-30.
13. Pagel O, Lorocho S, Sickmann A, Zahedi RP: **Current strategies and findings in clinically relevant post-translational modification-specific proteomics.***Expert Rev Proteomics* 2015, **12**:235-253.
14. Merrill AE, Coon JJ: **Quantifying proteomes and their post-translational modifications by stable isotope label-based mass spectrometry.***Curr Opin Chem Biol* 2013, **17**:779-786.
15. Sabari BR, Zhang D, Allis CD, Zhao Y: **Metabolic regulation of gene expression through histone acylations.***Nat Rev Mol Cell Biol* 2017, **18**:90-101.
16. Drazic A, Myklebust LM, Ree R, Arnesen T: **The world of protein acetylation.***Biochim Biophys Acta* 2016, **1864**:1372-1401.
17. Philp A, Rowland T, Perez-Schindler J, Schenk S: **Understanding the acetylome: translating targeted proteomics into meaningful physiology.***Am J Physiol Cell Physiol* 2014, **307**:C763-773.
18. Gil J, Ramirez-Torres A, Encarnacion-Guevara S: **Lysine acetylation and cancer: A proteomics perspective.***J Proteomics* 2017, **150**:297-309.
19. Filippakopoulos P, Knapp S: **Targeting bromodomains: epigenetic readers of lysine acetylation.***Nat Rev Drug Discov* 2014, **13**:337-356.
20. Choudhary C, Weinert BT, Nishida Y, Verdin E, Mann M: **The growing landscape of lysine acetylation links metabolism and cell signalling.***Nat Rev Mol Cell Biol* 2014, **15**:536-550.
21. Ouidir T, Kentache T, Hardouin J: **Protein lysine acetylation in bacteria: Current state of the art.***Proteomics* 2016, **16**:301-309.
22. Zhang Z, Tan M, Xie Z, Dai L, Chen Y, Zhao Y: **Identification of lysine succinylation as a new post-translational modification.***Nat Chem Biol* 2011, **7**:58-63.
23. Li X, Hu X, Wan Y, Xie G, Li X, Chen D, Cheng Z, Yi X, Liang S, Tan F: **Systematic identification of the lysine succinylation in the protozoan parasite *Toxoplasma gondii*.***J Proteome Res* 2014, **13**:6087-

24. Jin W, Wu F: **Proteome-Wide Identification of Lysine Succinylation in the Proteins of Tomato (*Solanum lycopersicum*)**. *PLoS One* 2016, **11**:e0147586.
25. Okanishi H, Kim K, Fukui K, Yano T, Kuramitsu S, Masui R: **Proteome-wide identification of lysine succinylation in thermophilic and mesophilic bacteria**. *Biochim Biophys Acta* 2017, **1865**:232-242.
26. Lin H, Su X, He B: **Protein lysine acylation and cysteine succination by intermediates of energy metabolism**. *ACS Chem Biol* 2012, **7**:947-960.
27. Shen C, Xue J, Sun T, Guo H, Zhang L, Meng Y, Wang H: **Succinyl-proteome profiling of a high taxol containing hybrid *Taxus* species (*Taxus x media*) revealed involvement of succinylation in multiple metabolic pathways**. *Sci Rep* 2016, **6**:21764.
28. Burkhardt JM, Gambaryan S, Watson SP, Jurk K, Walter U, Sickmann A, Heemskerk JW, Zahedi RP: **What can proteomics tell us about platelets?** *Circ Res* 2014, **114**:1204-1219.
29. Thomas S: **Platelets: handle with care**. *Transfus Med* 2016, **26**:330-338.
30. Kaisar M, van Dullemen LFA, Thezenas ML, Zeeshan Akhtar M, Huang H, Rendel S, Charles PD, Fischer R, Ploeg RJ, Kessler BM: **Plasma degradome affected by variable storage of human blood**. *Clin Proteomics* 2016, **13**:26.
31. Wong JW, McRedmond JP, Cagney G: **Activity profiling of platelets by chemical proteomics**. *Proteomics* 2009, **9**:40-50.
32. Sims C, Salliant N, Worth AJ, Parry R, Mesaros C, Blair IA, Snyder NW: **Metabolic tracing analysis reveals substrate-specific metabolic deficits in platelet storage lesion**. *Transfusion* 2017, **57**:2683-2689.
33. Smethurst PA: **Aging of platelets stored for transfusion**. *Platelets* 2016, **27**:526-534.
34. Navas-Carrillo D, Marin F, Valdes M, Orenes-Pinero E: **Deciphering acute coronary syndrome biomarkers: High-resolution proteomics in platelets, thrombi and microparticles**. *Crit Rev Clin Lab Sci* 2017, **54**:49-58.
35. D'Alessandro A, Kriebardis AG, Rinalducci S, Antonelou MH, Hansen KC, Papassideri IS, Zolla L: **An update on red blood cell storage lesions, as gleaned through biochemistry and omics technologies**. *Transfusion* 2015, **55**:205-219.
36. Capocelli KE, Dumont LJ: **Novel platelet storage conditions: additive solutions, gas, and cold**. *Curr Opin Hematol* 2014, **21**:491-496.
37. Heazlewood SY, Nilsson SK, Cartledge K, Be CL, Vinson A, Gel M, Haylock DN: **Progress in bio-manufacture of platelets for transfusion**. *Platelets* 2017, **28**:649-656.
38. Tyanova S, Temu T, Carlson A, Sinitsyn P, Mann M, Cox J: **Visualization of LC-MS/MS proteomics data in MaxQuant**. *Proteomics* 2015, **15**:1453-1456.
39. Ultsch A, Lotsch J: **Functional abstraction as a method to discover knowledge in gene ontologies**. *PLoS One* 2014, **9**:e90191.

40. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.***Nucleic Acids Res* 2000, **28**:27-30.
41. Du J, Yuan Z, Ma Z, Song J, Xie X, Chen Y: **KEGG-PATH: Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model.***Mol Biosyst* 2014, **10**:2441-2447.
42. Cox J, Mann M: **MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.***Nat Biotechnol* 2008, **26**:1367-1372.
43. Chou MF, Schwartz D: **Biological sequence motif discovery using motif-x.***Curr Protoc Bioinformatics* 2011, **Chapter 13**:Unit 13 15-24.
44. Cifani P, Kentsis A: **Towards comprehensive and quantitative proteomics for diagnosis and therapy of human disease.***Proteomics* 2017, **17**.
45. Gillet LC, Leitner A, Aebersold R: **Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing.***Annu Rev Anal Chem (Palo Alto Calif)* 2016, **9**:449-472.
46. Pagel O, Walter E, Jurk K, Zahedi RP: **Taking the stock of granule cargo: Platelet releasate proteomics.***Platelets* 2017, **28**:119-128.
47. Sotillo J, Doolan D, Loukas A: **Recent advances in proteomic applications for schistosomiasis research: potential clinical impact.***Expert Rev Proteomics* 2017, **14**:171-183.
48. Latorre A, Moscardo A: **Regulation of Platelet Function by Acetylation/Deacetylation Mechanisms.***Curr Med Chem* 2016, **23**:3966-3974.
49. Rauniyar N, Yates JR, 3rd: **Isobaric labeling-based relative quantification in shotgun proteomics.***J Proteome Res* 2014, **13**:5293-5309.
50. Keshishian H, Burgess MW, Specht H, Wallace L, Clauser KR, Gillette MA, Carr SA: **Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry.***Nat Protoc* 2017, **12**:1683-1701.
51. Zhu D, Hou L, Hu B, Zhao H, Sun J, Wang J, Meng X: **Crosstalk among proteome, acetylome and succinylome in colon cancer HCT116 cell treated with sodium dichloroacetate.***Sci Rep* 2016, **6**:37478.
52. Zhou H, Finkemeier I, Guan W, Tossounian MA, Wei B, Young D, Huang J, Messens J, Yang X, Zhu J, et al: **Oxidative stress-triggered interactions between the succinyl- and acetyl-proteomes of rice leaves.***Plant Cell Environ* 2017.
53. Zhen S, Deng X, Wang J, Zhu G, Cao H, Yuan L, Yan Y: **First Comprehensive Proteome Analyses of Lysine Acetylation and Succinylation in Seedling Leaves of Brachypodium distachyon L.***Sci Rep* 2016, **6**:31576.
54. Doll S, Burlingame AL: **Mass spectrometry-based detection and assignment of protein posttranslational modifications.***ACS Chem Biol* 2015, **10**:63-71.
55. Wang L, Xie R, Fan Z, Yang J, Liang W, Wu Q, Wu MX, Wang Z, Lu Y: **The contribution of oxidative stress to platelet senescence during storage.***Transfusion* 2019, **59**:2389-2402.

56. Lee S: **Post-translational modification of proteins in toxicological research: focus on lysine acylation.** *Toxicol Res* 2013, **29**:81-86.

Figures

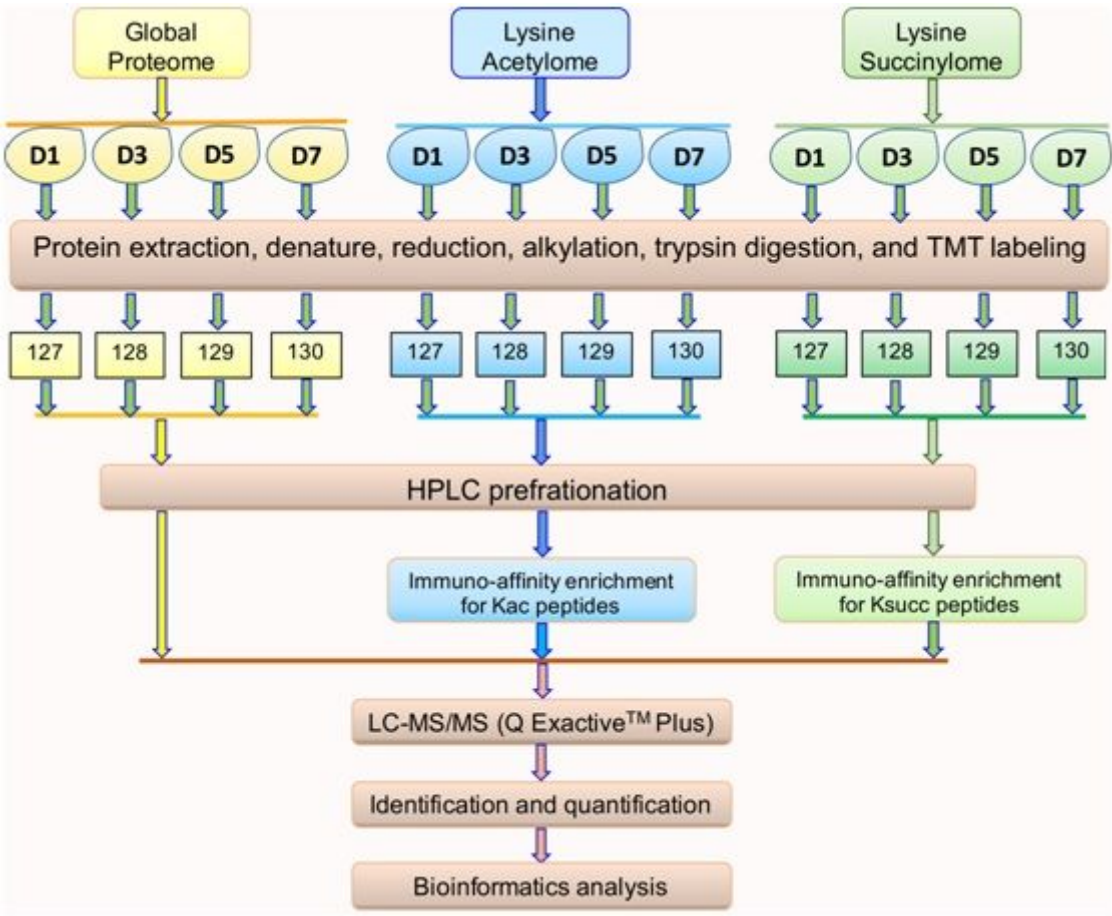


Figure 1

Workflow and strategy for quantitative multi-omics during platelet storage. The global proteome, lysine acetylome, and succinylome were performed using the same platelets in triplicate. There was no difference in preparing the platelet samples stored within the same storage time among the three proteomic groups. Note: D1 means platelet stored for 1 day.

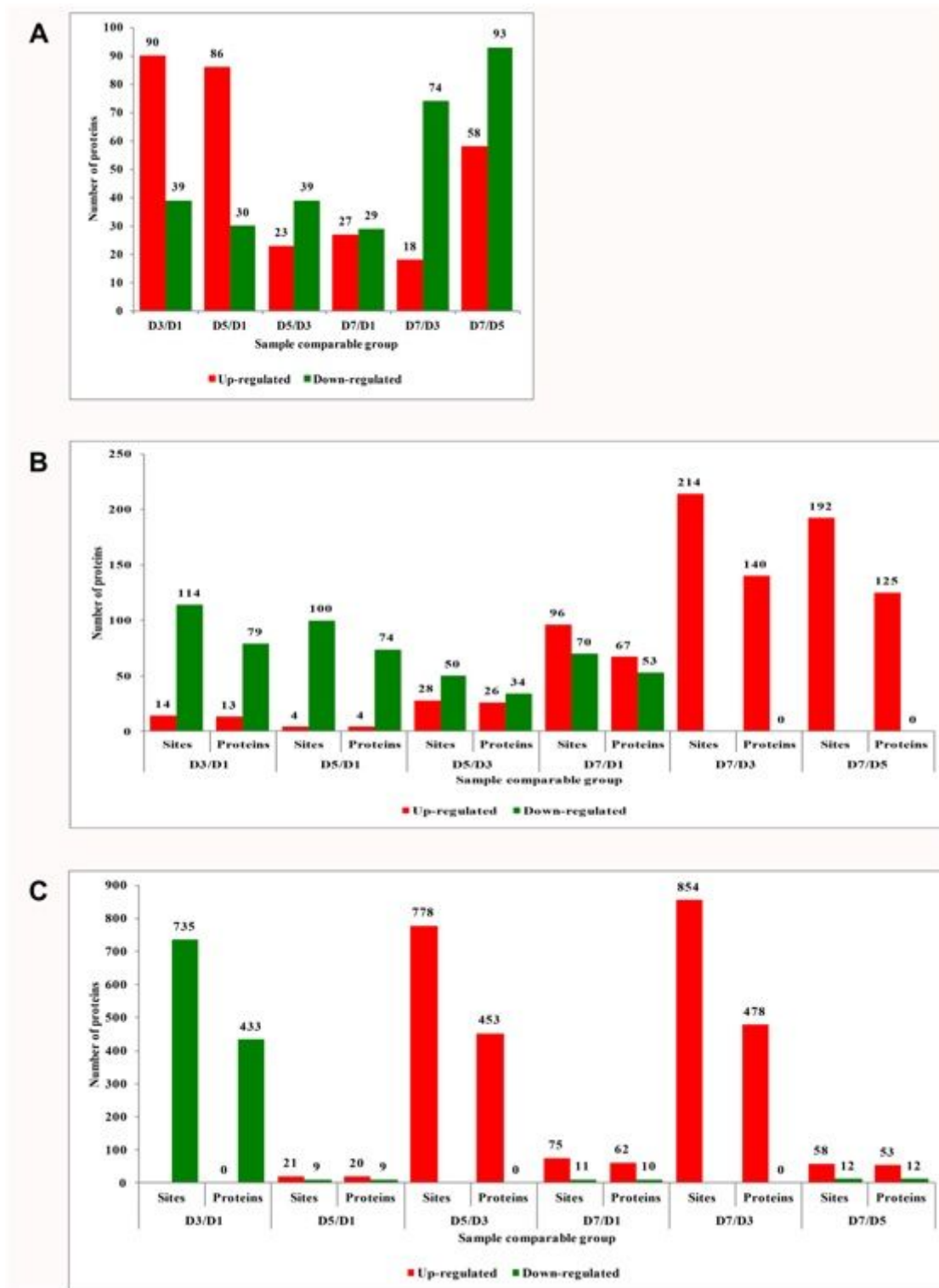


Figure 2

Differential expression of global proteins, Kac sites/proteins and Ksucc sites/proteins. A. up- and down-regulated proteins; B. up- and down-regulated Kac sites/proteins; C. up- and down-regulated Ksucc sites/proteins.

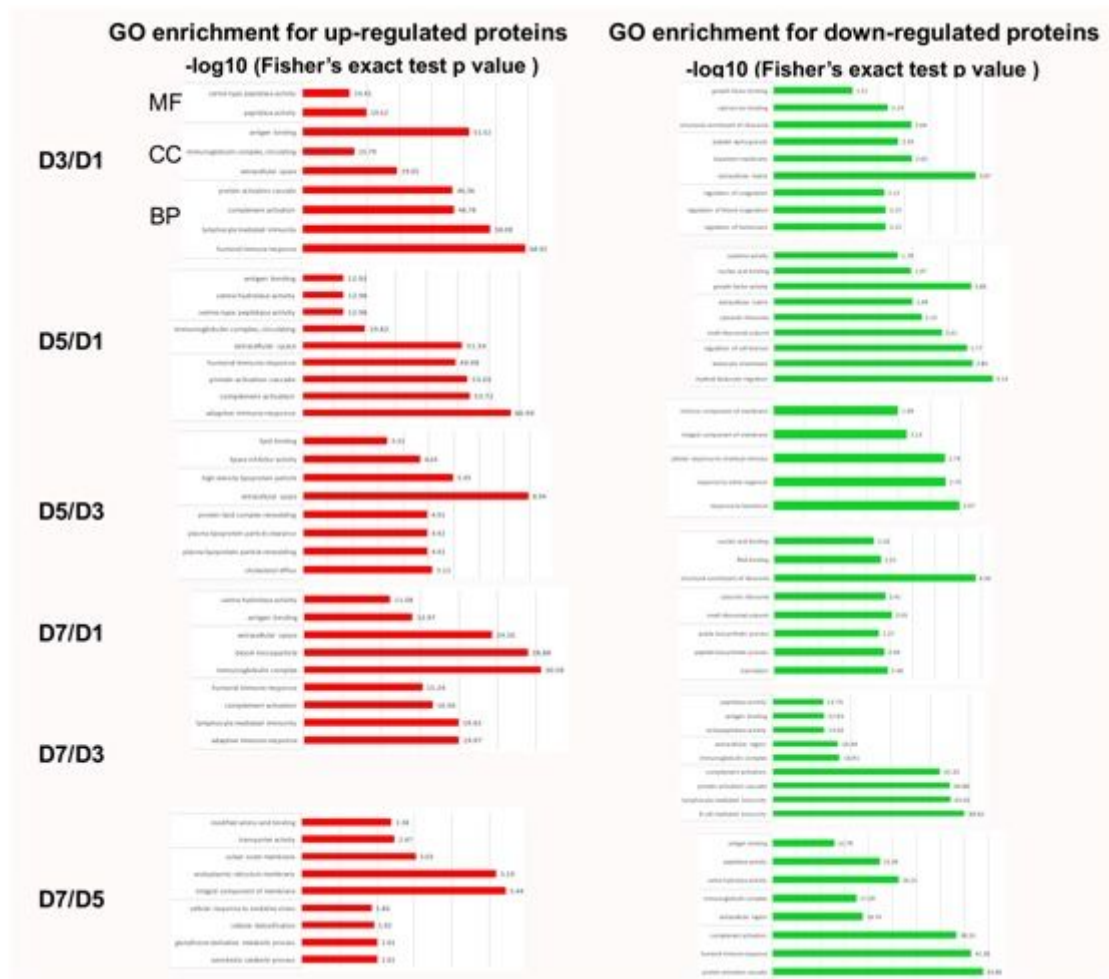


Figure 3

GO terms (MF, CC, and BP) for up-regulated and down-regulated proteins. Red indicates enrichment of up-regulated proteins, while Green indicates enrichment of down-regulated proteins. Note: MF: molecular function; CC: cellular component; BP: biological process.

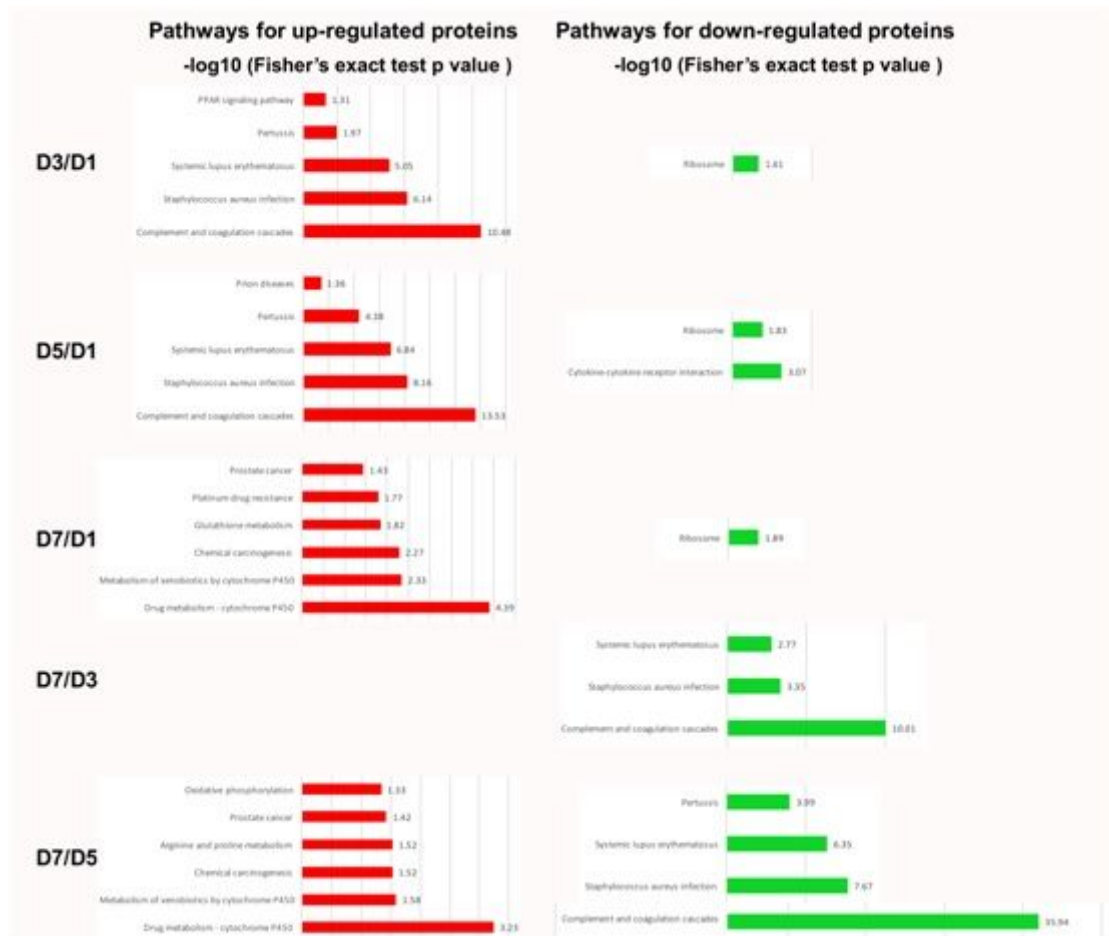


Figure 4

Pathway enrichment of differential proteins in specific comparison groups. Log-transformed Fisher's exact test was carried out to assess biochemical and physiological changes observed at different time points for up-regulated and down-regulated proteins. Note: Red indicates enrichment of up-regulated proteins. Green indicates enrichment of down-regulated proteins.

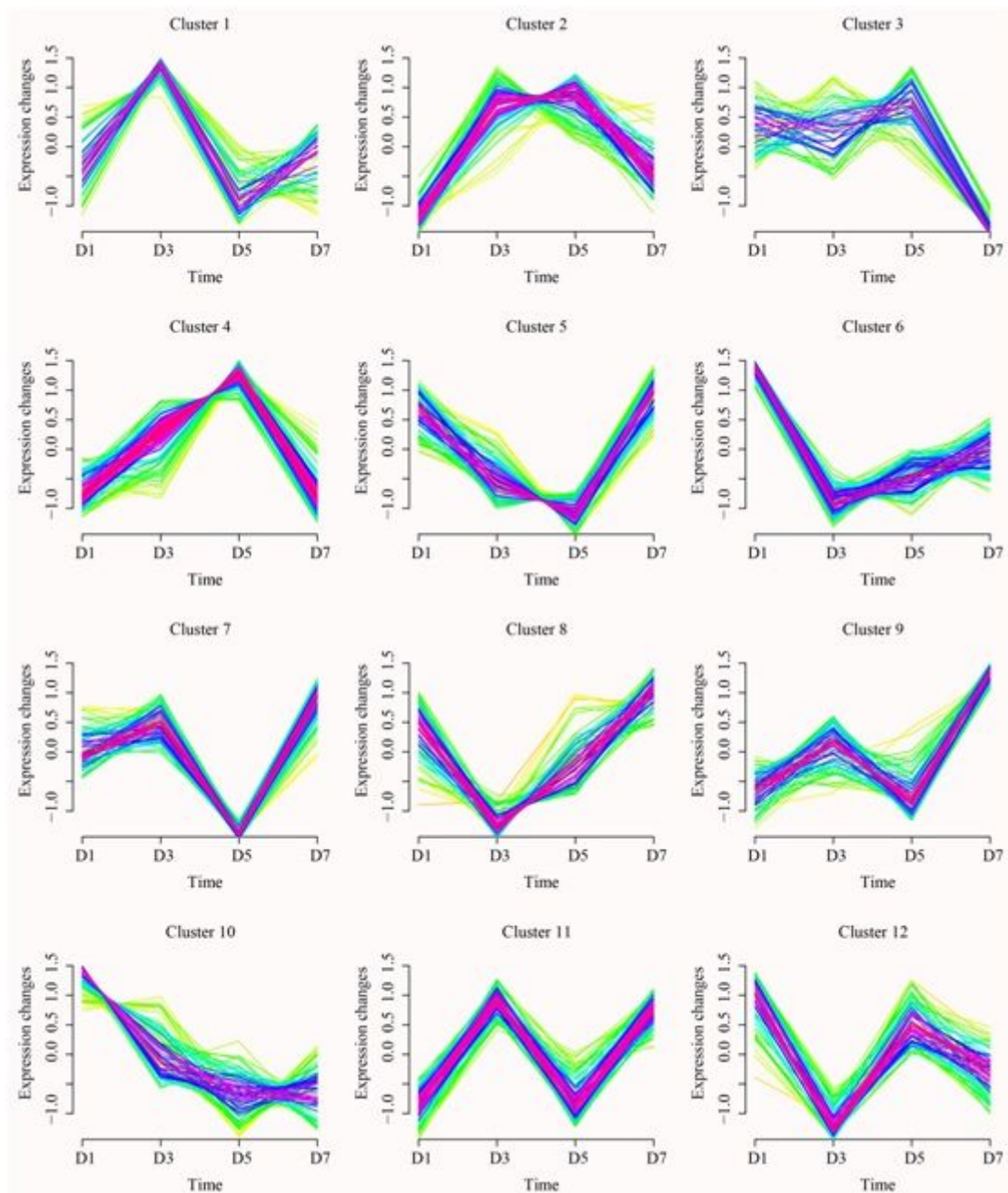


Figure 5

Enrichment-based clustering analysis of differential protein expression. 12 clusters of differential proteins with varied changes over storage time. Every cluster had specific traits among the groups, indicating differential patterns for the protein expressions of platelets among the 4 groups of storage time.

Motif analysis of all identified sites of Kac and Ksucc proteins. Gradient from green to red means the fold changes of the possibility of amino acid by log transformation. Red indicates the most likely amino acid to occur, while the green indicates the least likely. A. The motif analysis around the Kac sites. B. The motif analysis around the Ksucc sites.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supportinginformation.docx](#)