Int J Clin Exp Pathol 2016;9(7):7381-7392
www.ijcep.com /ISSN:1936-2625/IJCEP0025160

Original Article

Dynamic change of serum proteomics of occupational medicamentosa-like dermatitis induced by trichloroethylene

Wei Liu[1], Wenxu Hong[1], Yawen Zhang[3], Xiaohu Ren[1], Peimao Li[2], Zhimin Zhang[2], Haiyan Huang[1], Xinfeng Huang[1], Yanfang Zhang[2], Jianjun Liu[1]

[1]Key Laboratory of Modern Toxicology of Shenzhen, Shenzhen Center for Disease Control and Prevention, 8 Longyuan Road, Nanshan District, Shenzhen 518055, China; [2]Department of Clinical Laboratory of Shenzhen Hospital for Occupational Disease Treatment and Prevention, 70 Guiyuan North Road, Luohu District, Shenzhen 518001, China; [3]College of Life Science, Shenzhen University, Shenzhen 518060, China

Received January 28, 2016; Accepted May 20, 2016; Epub July 1, 2016; Published July 15, 2016

Abstract: Extensive researches about biomarkers of occupational medicamentosa-like dermatitis induced by Trichloroethylene (OMLDT) have been carried out in recent years. But dynamic change of protein biomarkers in serum has rarely been reported. The aim of our study was to explore the dynamic changing law of serum proteins/polypeptides in different periods of OMLDT, identify potential biomarkers and provide the scientific fundamentals for screening high-risk population and disease monitoring. We developed an approach in the combination of MB-WCX, MALDI-TOF-MS and ClinProTools software. Based on the alternations in the polypeptides fingerprint of serum (PFS), we built diagnostic models of OMLDT, and screened the special polypeptides biomarkers and further studied the dynamic changing law of different periods in typical OMLDT patients. We attained 72 peaks which were statistical content in OMLDT/Normal model, of which 52 peaks were differential peaks. We also obtained 69 significant peaks in OMLDT/TCE Contact model, but the differential peaks were 35. There were 21 specific peaks which were alike among the differential peaks between OMLDT/Normal model and OMLDT/TCE Contact model. Among the 21 specific peaks, we found 4 peaks, which were m/z 4109, 4267, 5065 and 9287 Da, changed nearly the same in 3 periods of 4 recurrent patients, and 2 peaks (4109 and 9173 Da) changed consistent in 3 periods of 3 stable patients. m/z 4109 Da may be the special serum biomarker of OMLDT. And Specific PFS may provide a new clue for clinical application.

Keywords: Occupational medicamentosa-like dermatitis induced by trichloroethylene, serum, proteomics, biomarkers, dynamic change

Introduction

Trichloroethylene (C_2HCl_3; TCE) is a major occupational hazard and environmental contaminant that can cause multisystem disorders such as occupational medicamentosa-like dermatitis. Occupational medicamentosa-like dermatitis induced by TCE (OMLDT) is dose-independent and potentially life threating disease. There is no obvious dose-effect relationship between the TCE exposure and incidence of OMLDT [1]. At present, the number of reported patients suffering from OMLDT in China has exceeded 500, and the mortality is about 9%-13% [2]. OMLDT has become an important occupational health issue in China. But the traditional protective measures including personal protective devices couldn't effectively reduce the incidence of OMLDT. One of the best ways to prevent the disease is to identify susceptible biomarkers used for screening employees before exposure and protecting susceptible populations.

At present, researches about OMLDT biomarker mainly focused on three aspects: gene polymorphism [3-7], serum proteomics [8-12] and TCE metabolites [13]. Haishan Li, et al. [3] identified genetic susceptible biomarkers associated with OMLDT in genes located in the human leukocyte antigen (HLA) region. Further research [4] had illustrated that HLA-B*13:01 was

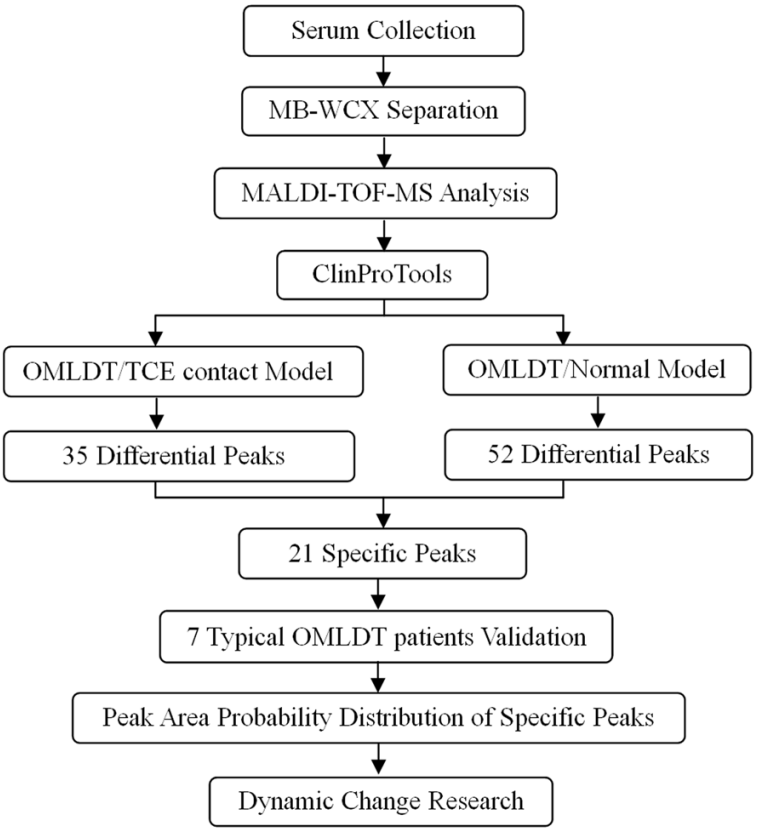


Figure 1. Schematic representation of study design. The collected sera from occupational medicamentosa-like dermatitis induced by trichloroethylene (OMLDT) patients, TCE contacts and healthy controls were purified by a weak cation exchange magnetic beads (MB-WCX) chromatography. The serum proteomic profiles were analyzed by matrix-assisted laser desorption/ionization-time of flight-mass spectrometry (MALDI-TOF-MS) and ClinProTools. After the statistical analysis by ClinProTools, twenty-one significantly different peaks which were alike between two models were selected and further validated in different periods of seven typical OMLDT patients. Those peaks were defined as specific peaks of OMLDT. Then we calculated the peak area probability of these specific peaks in each period and analyzed the dynamic change of peak area probability distribution in different periods of OMLDT patients. Finally, m/z 4109 Da characteristic peak may be the latent serum biomarker of OMLDT.

also identified in the patient of Japan. Later, they not only validated the association between HLA-B*13:01 allele and OMLDT, but also identified two new loci for the disease, one was on intron of the major histocompatibility complex class I chain related gene A (MICA), and another was between HLA-B and MICA, suggesting MICA was an important gene for the disease risk in addition to HLA-B*13:01. The combination of these single nucleotide polymorphisms (SNPs) could be an effective predicting biomarker for the disease among TCE-exposed populations [5]. Huang, et al. [12] compared the expression of proteins in the sera of OMLDT patients at three different stages using 2D-DIGE and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) strategy. They found that transthyretin (TTR), retinol binding protein 4 (RBP4) and haptoglobin could serve as potential serum biomarkers of OMLDT. Recently, Chloral hydrate (CH) seemed to be the culprit causative compound of OMLDT and the CH patch test could be potentially useful for the diagnosis of OMLDT and identification of subjects at risk of OMLDT [13].

Evidence existed that many disease processes were associated with quantitative and functional changes in the proteins of body fluids [14-16]. Serum was an easily accessible body fluid that contained different types of proteins released by various diseased tissues. Proteome analysis of the serum had played a central role in clinical diagnosis and monitoring and provided insight into disease pathophysiology and mechanisms. In our previous work [10, 11], we built disease model of OMLDT by MALDI-TOF MS and ClinProTools bioinformatics software and further identified two proteins, ATP-binding cassette transporter family A member 12 (ABCA12) and cationic trypsinogen (PRSS1), by a liquid chromatography (LC) coupled with tandem mass spectrometry (MS/MS) approach. These findings would be helpful for clinical diagnosis, and providing novel scientific clue for mechanisms studies of OMLDT. However, to date, the dynamic change of serum proteomics about OMLDT has rarely been explored. In the present study, to elucidate and understand the dynamic changing law in serum of different periods of OMLDT patients, we compared the alternations in the polypeptides fingerprint of serum (PFS) between

Table 1. The clinical information of the OMLDT, TCE contact and normal groups

Groups	Items	Training set	Test set	Total
OMLDT	Number	18	11	29
	Male/female	13/5	7/4	20/9
	Average age	24.06±8.81	23.45±7.22	23.75±8.01
TCE contact	Number	18	11	29
	Male/female	16/2	10/1	26/3
	Average age	28.44±5.84	24.36±5.07	26.40±5.45
Normal	Number	18	11	29
	Male/female	11/7	7/4	18/11
	Average age	23.72±4.44	22.73±2.57	23.22±3.50

Table 2. The clinical information of seven OMLDT patients for dynamic change research

Group	Patients no.	Age (year)	Sex	Liver damage
One	1	39	Female	Severe
	2	18	Male	Severe
	3	17	Male	Mild
	4	14	Male	Severe
Two	5	39	Female	Severe
	6	19	Male	Mild
	7	43	Male	Severe

OMLDT/Contact model and OMLDT/Normal model by the methods built in our previous researches. We further studied the peak areas probability distribution of specific peaks and validated in different periods of seven typical OMLDT patients. Finally, we found a characteristic peak (m/z 4109 Da) which may be the special serum biomarker of OMLDT (**Figure 1**). This study could appropriately reflect the dynamic change of serum proteins/polypeptides in different periods of OMLDT. And it may provide a new clue for clinical application.

Materials and methods

Clinical specimens

Eighteen OMLDT patients were enrolled in the study and 29 professional TCE contact people and 29 normal subjects were selected on an age-matched basis (±3 years). The samples were allocated to the training set and the test set for model establishment and validation, respectively. **Table 1** shows the basic clinical information of the three groups. The skin manifestations of the OMLDT patients enrolled in the study and the details for serum collection were described in our previous study [10, 11]. The clinical manifestations included fever, rash, liver damage and so on. Approximately 4 mL of blood was collected for each person and centrifuged at 4°C for 10 min to collect serum for MS analysis. Sera were collected by the Shenzhen hospital for Occupational Disease Treatment and Prevention after informed written consents were obtained. This study conformed to the ethical guidelines of the Declaration of Helsinki (World Medical Association, 1997) and was approved by the Institutional Review Board of this hospital.

Refer to the diagnostic criteria for occupational medicamentosa-like dermatitis induced by trichloroethylene (GBZ 185-2006) as defined by the Ministry of Health, the People's Republic of China, we further identified five periods of OMLDT: acute period, exfoliation period, bounce period, fade period and normal period or healed period according to patients condition, skin rashes characteristics and clinical examination indexes. Among eighteen OMLDT patients, seven patients with typical symptoms composed of 5 males and 2 females with a median age of 27 years (range 14-43 years) were used for dynamic change research. They were divided into two groups (**Table 2**). Group one contained four patients, whose condition were recurrent. The sera of exfoliation period, bounce period and fade period were collected and analyzed. Group two included three stable patients, who executed the criteria that the symptoms were in line with changes in the basic process of OMLDT, such as acute period, exfoliation period and healed period. The sera of these three periods were collected and analyzed.

Mass spectrometry and data interpretation

Peptidome separation of the serum samples and linear MALDI-TOF-MS analysis were described in our previous work [10, 11]. Spectra had been collected automatically using the AutoXecute™ software (Bruker Daltonics) for fuzzy-controlled adjustment of critical instrument settings to generate raw data of optimized quality. The ClinProTools™ software (Version 2.2; Bruker Daltonics) had been used for

all data interpretation steps, which started with a raw data pretreatment, including normalization of a set of spectra derived from a patient cohort, internal signal alignment using prominent internal signal peaks and a peak picking procedure. The signal-to-noise (S/N) ratio was higher than 5.0. Peak calculation was used areas. The whole data pretreatment has been completed using default settings and was performed automatically, without any user interaction.

ClinProTools offered three basic workflows: Peak Statistic Calculation, Model Generation and Classification. Peak Statistic Calculation included spectra recalibration and average spectra calculation, peak picking and peak calculation as well as peak statistic calculation. Model Generation included spectra recalibration and average spectra calculation, peak calculation and model generation based on the selected classification algorithm. Classification can be used to quickly classify test spectra with an existing model. This workflow included selecting the model to use and the spectra to classify, data preparation of these spectra according to the settings saved in the model and their classification. In this study, differential peaks were selected on the basis of statistical differences between relative protein/peptide ion peak areas with OMLDT and professional TCE contacts or healthy controls. These calculations were performed independently for peak areas. The difference of peak average areas (DAve) was larger; the separate power was stronger. The generally accepted value was more than 10. To characterize the peptidomes of OMLDT patients, we focused our attention on peaks where the DAve was larger than 11. The pretreated data were used to build models that could discriminate OMLDT patients from TCE contacts or normal subjects by a Supervised Neural Network (SNN) Algorithm. Model-building was performed by selecting a small subset of relevant peaks and establishing clusters using the peak areas. The accuracy of models were evaluated and validated by the test data.

Dynamic change research of serum proteomics for OMLDT

Through analyzing the differential peaks between the acute/contact model and the acute/normal model by the statistic methods, we figure out the specific peaks of OMLDT. To understand the dynamic changing law of different periods in the whole course and screen the special proteins/polypeptides markers, we analyzed the dynamic change of PFS in different periods of 7 typical patients according to the specific peaks. The peak areas probability distribution could effectively reflect the expression level of serum proteins/polypeptides in some period. Single peak area divided by all peak areas was the peak areas probability. The dynamic change of serum proteins/polypeptides expression level could be observed by comparing the peak areas probability distribution in different periods. We calculated the peak areas probability distribution of specific peaks in different periods of each patient and validated in different periods of seven typical OMLDT patients.

Statistical analysis

ClinProTools supports various statistical tests and methods, including t-test, ANOVA test, Wilcoxon test, Kruskal-Wallis test and Anderson-Darling test. The Anderson-Darling (AD) test in the case of ClinProTools has been adapted to test for normal distributions. Each of these tests calculates the so-called *P*-value, which is the probability that an observed effect is simply due to chance. Accordingly, the lower the *P* value, the better a respective peak signal is suited to be used to separate the two classes. The generally accepted limit of the *P* values to consider a result significant is defined as 0.05 (0.01 for highly significant results). Hence, at first we should evaluate the *P*-value for the AD test. If it is above e.g. 0.05, we should consider the t-test or ANOVA test; otherwise, the result from Wilcoxon/Kruskal-Wallis (W/KW) test has to be evaluated. These calculations were done independently for peak areas.

Results

OMLDT/normal model

ClinProTools provides a list of peaks sorted according to the statistical significance to differentiate between both classes. As a result, we got 158 peaks in OMLDT/Normal model, of which 72 peaks were statistical content (Supplementary Figures 1, 2). There were 52 differential peaks which were significantly different in the 72 peaks (**Table 3**). Among these,

Table 3. Fifty-two differential peaks in sera from OMLDT patients

Mass	DAve	Peak Average Areas		PTTA	PWKW	PAD
		OMLDT	Normal			
Low-expressed proteins/polypeptides in OMLDT group						
1076.70	13.02	16.59±15.48	29.62±25.28	0.132	0.0258	< 0.000001
2046.76	13.34	19.21±4.31	32.54±5.26	< 0.000001	< 0.000001	0.2
2082.87	21.72	15.65±7.44	37.37±41.35	0.0853	0.0422	< 0.000001
2106.30	59.87	43.48±15.11	103.36±16.09	< 0.000001	< 0.000001	0.0105
2125.67	12.12	12.91±4.84	25.03±14.22	0.00813	< 0.0000368	< 0.000001
2134.50	18.09	12.86±4.47	30.95±6.01	< 0.000001	< 0.000001	0.0239
2670.34	15.72	25.52±8.64	41.24±12.39	0.000645	0.000227	0.0676
2933.36	11.12	11.85±4.13	22.97±4.55	< 0.000001	< 0.0000038	0.0613
2953.79[$]	23.36	18.43±9.17	41.79±12.19	< 0.0000049	< 0.0000043	0.0139
3263.67[$]	18.87	12.79±4.98	31.67±9.60	< 0.0000016	< 0.000001	0.056
3275.77	30.47	21.34±6.95	51.81±50.79	0.0508	0.00037	< 0.000001
3335.18	17.05	24.30±8.24	41.36±12.26	0.000237	0.000096	0.0491
4054.72	49.01	184.69±58.72	233.69±55.29	0.0357	0.0491	0.942
4091.68	85.46	168.98±52.54	254.43±65	0.000688	0.000286	0.162
4109.94[$]	24.32	32.00±8.01	56.32±11.98	< 0.0000014	< 0.000002	0.0651
4123.84	36.31	69.65±16.42	105.95±19.67	< 0.0000148	< 0.0000104	0.824
4146.26	14.78	32.57±8.03	47.35±8.06	< 0.0000448	< 0.0000397	0.877
4168.84[$]	30.60	54.11±17.06	84.71±18.58	0.000111	< 0.0000368	0.648
4195.40	50.52	77.07±29.10	127.59±32.04	0.00016	< 0.0000469	0.584
4210.66	431.66	480.52±224.01	912.18±257.10	< 0.0000675	< 0.0000368	0.739
4248.33[$]	68.01	49.64±18.46	117.65±23.23	< 0.000001	< 0.000001	0.168
4267.30[$]	109.50	62.35±27.70	171.85±38.09	< 0.000001	< 0.000001	0.03
4283.87[$]	32.44	26.11±17.29	58.55±34.28	0.00509	< 0.0000397	< 0.0000178
5065.37[$]	21.54	15.53±5.31	37.07±13.91	< 0.0000448	< 0.0000012	0.00682
5337.17	119.05	103.40±55.40	222.45±94.67	0.000519	0.000286	0.226
5374.61	16.15	12.78±5.28	28.93±11.48	0.000128	< 0.0000038	0.0491
5904.68[$]	95.12	55.70±48.94	150.82±59.73	0.000102	< 0.0000158	0.00314
5939.33	11.30	19.77±6.18	31.07±9.65	0.0012	0.00033	0.00799
High-expressed proteins/polypeptides in OMLDT group						
1001.83[$]	13.62	40.50±18.95	26.88±5.47	0.0231	0.0455	0.000389
1526.62[$]	18.56	29.33±18.97	10.77±2.60	0.0029	< 0.000001	< 0.000001
1761.95[$]	27.87	55.60±30.19	27.73±7.46	0.00451	0.000262	< 0.000001
1787.48[$]	15.73	43.49±18.18	27.76±6.68	0.00761	0.00822	0.000524
2483.27	16.27	30.02±15.81	13.76±3.29	0.00196	0.00013	< 0.000006
2546.17	13.21	24.80±22.19	11.59±20	0.0512	0.000203	< 0.000001
2732.15	13.82	27.66±16.89	13.84±4.48	0.0096	0.0304	< 0.000002
3883.92	19.70	61.84±23.29	42.14±15.19	0.0155	0.0204	0.267
4964.34	195.11	235.36±150.04	40.25±13.81	0.000267	< 0.0000043	< 0.0000031
5001.95	23.15	36.12±23.70	12.97±5.54	0.00297	0.000227	< 0.0000098
7764.85	195.81	369.62±182.07	173.81±78.93	0.00162	0.000497	0.0214
7802.47	16.52	40.27±17.59	23.75±9.78	0.00597	0.00312	0.087
7922.44	17.63	45.57±19.04	27.94±6.70	0.00477	0.00312	0.000859
8140.86	26.03	53.28±21.01	27.25±5.73	0.00043	< 0.0000397	0.000951
8564.50	14.28	31.34±36.50	17.06±2.73	0.189	0.00341	< 0.000001
8861.06[$]	14.08	48.52±16.06	34.45±6.86	0.00761	0.00599	0.00168

8915.65$	14.56	28.18±12.13	13.62±2.77	0.000521	< 0.0000675	< 0.0000652
8930.74$	12.84	20.42±10.83	7.57±1.43	0.000573	0.000151	< 0.0000012
9060.82$	17.87	53.35±18.44	35.48±9.10	0.00427	0.00341	0.00204
9131.79$	55.35	97.94±51.48	42.59±10.62	0.00135	0.00037	< 0.0000039
9173.97$	35.21	70.71±28.76	35.50±11.96	0.000505	0.000773	< 0.0000441
9287.18$	559.73	1654.22±651.91	1094.49±551.54	0.024	0.0222	0.0337
9420.63$	70.17	135.95±73.37	65.78±18.80	0.00357	0.00113	< 0.0000295
11729.45	11.37	28.96±10.37	17.59±2.47	0.0012	< 0.0000087	0.000243

Dave, Difference of peak average areas; PTTA, *P* value of t-test or ANOVA test; PWKW, *P* value of Wilcoxon test or Kruskal-Wallis test; PAD, *P* value of Anderson-Darling test. $, specific peaks of OMLDT for dynamic change research.

Table 4. Model Establishment and Validation calculated by SNN algorithm

Model	Establishment		Validation	
	Cross validation	Recognition capability	Sensitivity	Specificity
OMLDT/Normal	94.44	100	100	66.7
OMLDT/TCE Contact	97.22	100	81.3	100

28 peaks were low expressed and 24 peaks were high expressed in OMLDT sera compared with those in healthy controls. We built the model by SNN algorithm. Three peptide ion signatures (m/z 2106.3, 2134.5 and 3263.67) were provided as class predictors for a cross-validation set, discriminating OMLDT samples from healthy controls (Supplementary Figures 3, 4). Its cross validation and recognition capability were 94.44% and 100%, respectively. And the sensitivity and specificity were 100% and 66.7%, respectively (**Table 4**).

OMLDT/TCE contact model

We attained 157 peaks in the training data set of OMLDT/TCE Contact model, of which 69 peaks were statistically significant (Supplementary Figures 1, 2). There were 35 differential peaks in the 69 peaks (**Table 5**). Among the 35 peaks, 13 peaks were down-regulated and 22 peaks were up-regulated in OMLDT groups compared with those in TCE Contact controls. We also built the model by SNN algorithm, which consisted of 1450.33, 1866.16, 3262.39, 4109.55, 5064.85, 5248.05, 5956.57 and 6667.04 Da. Its cross validation and recognition capability were 97.22% and 100%, respectively. And the sensitivity and specificity were 81.3% and 100%, respectively, which explained good separating capacity (**Table 4**).

Results of cpecific peaks

To find some specific signals for OMLDT, we compared the differential peaks between the OMLDT/Normal model and the OMLDT/TCE contact model (Supplementary Figures 1, 2, 3, 4). Twenty-one peaks were found in both models and the change of expression level was consistent (**Tables 3** and **5**). They were likely to be the specific peaks of OMLDT. Among them, nine were lower expressed and twelve were high expressed in the comparison of TCE contact group and Normal group. These specific peaks could be especially well visualized in the mass region among 1000 to 1800 Da, 2900 to 4300 Da, 5000 to 6000 Da, and 8800 to 9500 Da.

Dynamic change research for OMLDT

OMLDT specific peaks were further selected for analyzing and validating in seven typical cases. We studied the dynamic change of specific peaks in different periods by calculating peak area probability distribution. Seven patients were divided into two groups: recurrent group (group one, four cases) and non-recurrent group or stable group (group two, three cases) according to the disease situation. The dynamic change law of specific peaks in two groups was observed respectively.

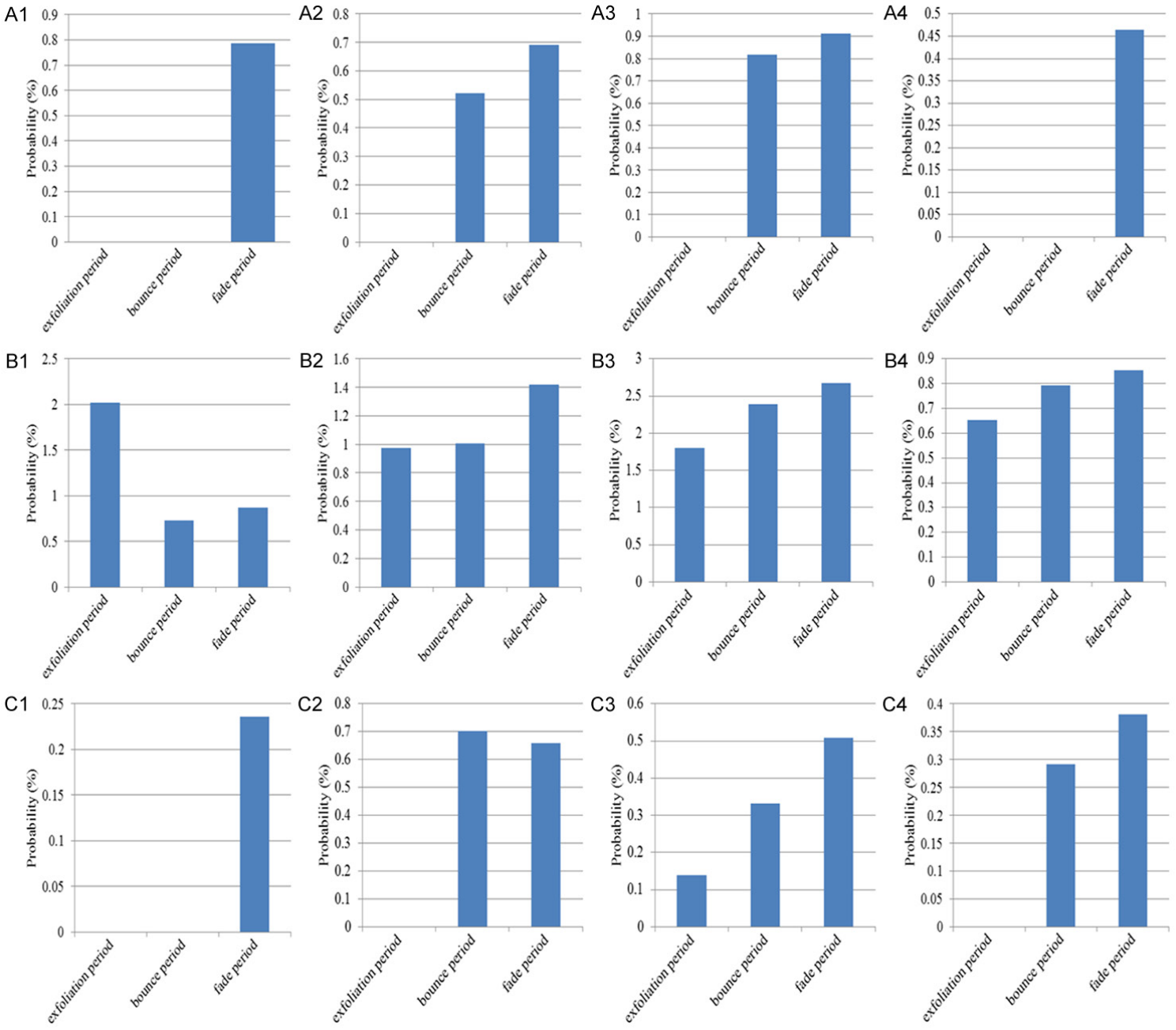
Among 21 specific peaks, we found 4 peaks, which were m/z 4109 Da, 4267 Da, 5065 Da and 9287 Da, changed nearly the same in 3 periods of 4 recurrent patients (**Figure 2A-D**). The expression change of m/z 4109 Da and 5065 Da presented the overall upward trend. The expression level of m/z 4267 Da presented the upward trend in No.2, 3 and 4 whereas

Table 5. Thirty-five differential peaks in sera from OMLDT patients

Mass	DAve	Peak Average Areas		PTTA	PWKW	PAD
		OMLDT	TCE Contact			
Low-expressed proteins/polypeptides in OMLDT group						
1450.33	18.98	13.53±4.97	32.52±9.92	0.00000685	< 0.000001	0.0578
3262.39$	13.64	10.91±4.02	24.56±7.09	0.00000685	< 0.000001	0.0914
5064.85$	18.66	15.21±5.26	33.87±9.62	0.00000685	0.00000135	0.0136
4109.55$	20.30	33.05±8.17	53.35±12.86	0.000133	0.0000313	0.0543
1866.16	30.26	13.37±6.44	43.63±20.69	0.000212	< 0.000001	0.000361
4266.61$	48.09	61.62±27.57	109.71±35.25	0.000815	0.000195	0.106
5904.28$	65.66	54.15±47.83	119.81±49.17	0.00237	0.000456	0.0625
4247.85$	27.39	47.90±17.80	75.29±24.06	0.0034	0.00361	0.507
2953.30$	11.08	17.37±8.90	28.45±10.14	0.00666	0.000608	0.00435
2660.95	90.65	88.69±72.32	179.34±96.64	0.0108	0.00263	0.0311
4168.15$	13.57	53.02±16.65	66.60±16.15	0.0445	0.0302	0.476
4283.20$	22.74	27.23±17.66	49.97±56.49	0.177	0.00645	< 0.000001
1981.38	16.03	15.58±2.75	31.61±45.03	0.208	0.0365	< 0.000001
High-expressed proteins/polypeptides in OMLDT group						
4680.29	16.39	45.72±11.81	29.34±7.45	0.000482	0.0000862	0.121
6667.04	61.65	103.66±47.07	42.01±14.51	0.000514	0.0000173	0.000427
8915.89$	11.95	22.24±9.63	10.29±2.06	0.000815	0.000163	0.0000765
9173.82$	34.26	67.89±27.83	33.63±10.92	0.000815	0.000972	0.000479
9131.96$	56.03	90.99±48.38	34.96±10.63	0.00129	0.000172	0.00000892
8930.55$	11.98	20.94±11.05	8.96±1.62	0.00237	0.000358	0.00000193
6630.29	224.54	335.41±210.60	110.87±43.14	0.00253	0.000972	0.0000048
1761.86$	24.44	41.68±23.95	17.24±3.48	0.00329	0.0000688	< 0.000001
1040.50	31.83	70.87±31.20	39.04±10.74	0.00341	0.00361	0.0000146
1786.76$	12.75	27.31±12.69	14.56±3.13	0.00348	0.00237	0.0000311
6432.17	50.32	91.68±50.46	41.37±15.85	0.0038	0.000944	0.0000149
1526.55$	18.15	30.28±19.18	12.12±3.28	0.00516	< 0.000001	< 0.000001
1015.15	19.51	44.39±20.77	24.88±7.52	0.00587	0.00489	0.0000766
9060.91$	18.27	54.73±18.75	36.47±12.53	0.00778	0.00513	0.000754
9420.36$	63.96	137.43±74.50	73.47±18.18	0.00882	0.00513	0.000215
4711.90	16.10	42.12±19.04	26.03±3.49	0.00888	0.0000767	0.00000193
9287.22$	653.26	1657.78±653.97	1004.52±524.09	0.00888	0.00645	0.0651
4644.86	118.46	366.11±95.71	247.65±119.41	0.00897	0.00712	0.749
3316.54	25.95	56.63±30.62	30.67±8.98	0.00907	0.0117	0.0000843
6587.98	21.00	46.79±25.18	25.79±9.60	0.0108	0.0117	0.00000111
1001.60$	18.82	51.46±22.92	32.65±12.09	0.0156	0.0151	0.00599
8860.93$	13.61	49.19±15.51	35.58±11.44	0.0167	0.0117	0.00999

Dave, Difference of peak average areas; PTTA, *P* value of t-test or ANOVA test; PWKW, *P* value of Wilcoxon test or Kruskal-Wallis test; PAD, *P* value of Anderson-Darling test. $, specific peaks of OMLDT for dynamic change research.

downward trend in No.1 patient. The expression level of m/z 9287 Da presented an increasing trend in bounce period and then decreasing tendency in fade period of No.1, 3 and 4. But it was overall downward trend in No.2 patient. We then found the expression change rule of two specific peaks (m/z 4109 Da and 9173 Da) in three OMLDT cases among the acute period, exfoliation period and recovery period was exactly the same (**Figure 2E** and **2F**). The expression levels of m/z 4109 Da showed an increasing trend, but the expression levels of


A1
A2
A3
A4
B1
B2
B3
B4
C1
C2
C3
C4
Probability (%)
exfoliation period
bounce period
fade period

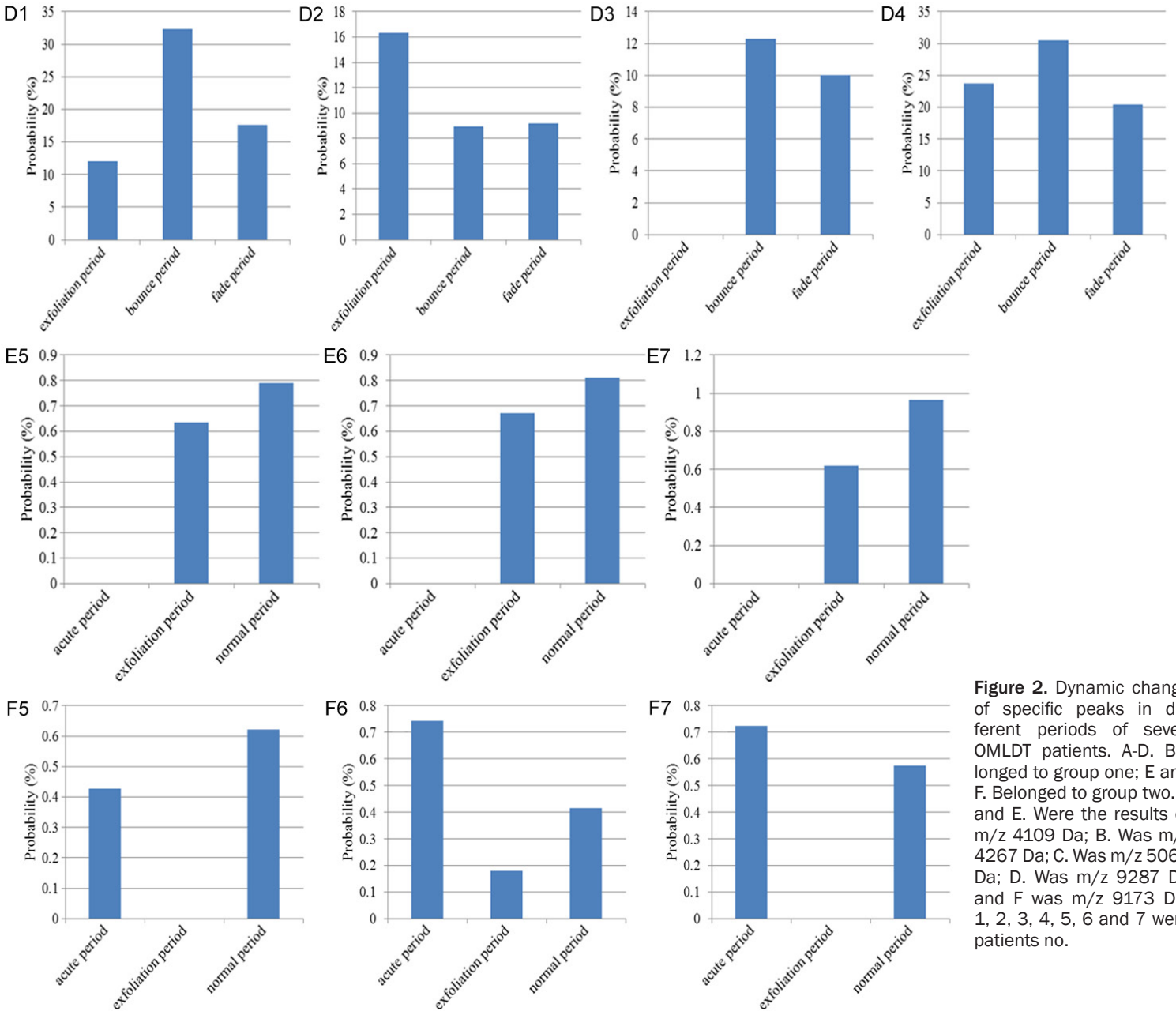


Figure 2. Dynamic change of specific peaks in different periods of seven OMLDT patients. A-D. Belonged to group one; E and F. Belonged to group two. A and E. Were the results of m/z 4109 Da; B. Was m/z 4267 Da; C. Was m/z 5065 Da; D. Was m/z 9287 Da and F was m/z 9173 Da. 1, 2, 3, 4, 5, 6 and 7 were patients no.

m/z 9173 Da appeared a decreasing trend in exfoliation phase and then an increased trend in the recovery phase. Interestingly, m/z 4109 Da appeared in both groups, and the expression level was increased with the course of disease. So it may be the special serum biomarker of OMLDT.

Discussion

Mass spectrometry-based search for biomarker patterns was widely recognized as a valuable research tool for predictive medicine and pharmacological monitoring [17-20]. ClinProTools was the data interpretation software of the mass spectrometry-based ClinProt™ solutions for biomarker analysis [21]. The ClinProTools software was used for the data interpretation of MALDI-TOF spectra derived from serum samples of different groups. In this study, we had used two different control groups, namely TCE contact controls and healthy controls, which enabled us to distinguish general disease marker candidates from those that were specific for the OMLDT disease. We found that sera from OMLDT patients can be distinguished from healthy controls based on an array of 52 discriminated proteins/peptides. Most of them were consistent with our previous work [10]. On this basis, we further compared the differential peaks between OMLDT and TCE contact model. Thirty-five peaks were found and also in accordance with our preliminary study [11].

Within clinical proteomics, we were in general confronted with a small set of samples and a large number of identified peaks. The identification of biomarker candidates within ClinProTools focuses on the detected peaks over a given set of spectra. Peak area probability is a good indicator and rarely report in researches. The area of a peak is obtained by integrating the intensities over the region of the peak according to the full intensity values. Peak ar eas have a smaller variation between spectra than intensities of single points have. In our work, we calculated the peak area probability of thirty-five specific peaks in different periods of seven OMLDT patients and observed the change trend to find out the potential law. Because of differences in individual cases or may be complexity of patients condition changes [13, 22, 23], peak area probability distribution most of specific peaks had no obvious law. Among 21 specific peaks, we found 4 peaks, which were 4109, 4267, 5065 and 9287 Da, changed nearly the same in 3 periods of 4 recurrent patients, and 2 peaks (4109 and 9173 Da) changed consistent in 3 periods of 3 stable patients. We were exciting to discover m/z 4109 Da peak appeared both in bounce group and non-bouncing group, and its expression level changed with the disease progression showed increasing tendency. Therefore, we initially considered m/z 4109 Da specific peak may be OMLDT unique biomarker.

To date, however, there are still no available biomarkers for predicting who is at high risk of developing occupational medicamentosa-like dermatitis among TCE-exposed workers in clinical application. There were also some limitations and disadvantages in our study. Firstly, the sample size was not enough. A larger data set was needed for such a confirmation. Secondly, potential biomarker peaks should be selected for in-depth analysis after further enrichment to reduce the sample complexity and identified by tandem mass spectrometry (TOF/TOF MS). Moreover, identification via TOF/TOF fragment analysis should be performed to provide biological relevance to the statistical analysis.

Conclusively, ClinProTools was a very powerful software tool for the data interpretation in the field of clinical proteomics. It combined efficient visualization with automated data pretreatment and intuitive statistical analysis. In the present study, we had demonstrated that disease-specific proteomic fingerprints were present in the sera of OMLDT patients. Peak area probability distribution could well reflect the dynamic change law of polypeptides fingerprint in serum in the development of disease. These results provided scientific clues for further identification and crowd verification.

Acknowledgements

This work was supported by the NSFC (the National Natural Science Foundation of China) [81273126], the project of Guangdong science and technology plan [2013B031800001], the Project of Shenzhen Basic Research Plan [JCYJ-20130329103949656,JCYJ2015040210213-5497] and the Future Industry Special Project of Shenzhen (ZDSYS20140509101335476).

Disclosure of conflict of interest

None.

Address correspondence to: Jianjun Liu, Key Laboratory of Modern Toxicology of Shenzhen, Shenzhen Center for Disease Control and Prevention, 8 Longyuan Road, Nanshan District, Shenzhen 518055, China. Tel: +86 0755 2550 8584; Fax: +86 0755 2550 8584; E-mail: bio-research@hotmail.com; junii8@126.com; Yanfang Zhang, Department of Clinical Laboratory of Shenzhen Hospital for Occupational Disease Treatment and Prevention, 70 Guiyuan North Road, Luohu District, Shenzhen 518001, China. Tel: +86 0755 61385573; E-mail: 13923431511@139.com

References

[1] Vineis P, Ahsan H, Parker M. Genetic screening and occupational and environmental exposures. Occup Environ Med 2005; 62: 657-62, 597.

[2] Kamijima M, Hisanaga N, Wang H, Nakajima T. Occupational trichloroethylene exposure as a cause of idiosyncratic generalized skin disorders and accompanying hepatitis similar to drug hypersensitivities. Int Arch Occup Environ Health 2007; 80: 357-70.

[3] Li H, Dai Y, Huang H, Li L, Leng S, Cheng J, Niu Y, Duan H, Liu Q, Zhang X, Huang X, Xie J, Feng Z, Wang J, He J, Zheng Y. HLA-B*1301 as a biomarker for genetic susceptibility to hypersensitivity dermatitis induced by trichloroethylene among workers in China. Environ Health Perspect 2007; 115: 1553-6.

[4] Watanabe H, Tohyama M, Kamijima M, Nakajima T, Yoshida T, Hashimoto K, Iijima M. Occupational trichloroethylene hypersensitivity syndrome with human herpesvirus-6 and cytomegalovirus reactivation. Dermatology 2010; 221: 17-22.

[5] Dai Y, Chen Y, Huang H, Zhou W, Niu Y, Zhang M, Bin P, Dong H, Jia Q, Huang J, Yi J, Liao Q, Li H, Teng Y, Zang D, Zhai Q, Duan H, Shen J, He J, Meng T, Sha Y, Shen M, Ye M, Jia X, Xiang Y, Huang H, Wu Q, Shi M, Huang X, Yang H, Luo L, Li S, Li L, Zhao J, Li L, Wang J, Zheng Y. Performance of genetic risk factors in prediction of trichloroethylene induced hypersensitivity syndrome. Sci Rep 2015; 5: 12169.

[6] Dai Y, Leng S, Li L, Niu Y, Huang H, Cheng J, Zheng Y. Genetic polymorphisms of cytokine genes and risk for trichloroethylene-induced severe generalized dermatitis: a case-control study. Biomarkers 2004; 9: 470-78.

[7] Dai Y, Leng S, Li L, Niu Y, Huang H, Liu Q, Duan H, Cheng J, Liu Q, Zheng Y. Effects of genetic polymorphisms of N-Acetyltransferase on trichloroethylene-induced hypersensitivity dermatitis among exposed workers. Ind Health 2009; 47: 479-86.

[8] Liu J, Xing X, Huang H, Jiang Y, He H, Xu X, Yuan J, Zhou L, Yang L, Zhuang Z. Identification of antigenic proteins associated with trichloroethylene-induced autoimmune disease by serological proteome analysis. Toxicol Appl Pharmacol 2009; 240: 393-400.

[9] Huang Z, Yue F, Yang X, Xia L, Chen C, Qiu X, Huang J, Li L, Kamijima M, Nakajima T, Huang H. Upregulation of calprotectin and downregulation of retinol binding protein in the serum of workers with trichloroethylene-induced hypersensitivity dermatitis. J Occup Health 2012; 54: 299-309.

[10] Hong WX, Liu W, Zhang Y, Huang P, Yang X, Ren X, Ye J, Huang H, Tang H, Zhou G, Huang X, Zhuang Z, Liu J. Identification of serum biomarkers for occupational medicamentosa-like dermatitis induced by trichloroethylene using mass spectrometry. Toxicol Appl Pharmacol 2013; 273: 121-9.

[11] Liu W, Hong WX, Zhang Y, Huang P, Yang X, Ren X, Huang H, Liu J. Proteomic profiling of occupational medicamentosa-like dermatitis induced by trichloroethylene in serum based on MALDI-TOF MS. Clin Exp Med 2015; 15: 519-526.

[12] Huang P, Ren X, Huang Z, Yang X, Hong W, Zhang Y, Zhang H, Liu W, Huang H, Huang X, Wu D, Yang L, Tang H, Zhou L, Li X, Liu J. Serum proteomic analysis reveals potential serum biomarkers for occupational medicamentosa-like dermatitiscaused by trichloroethylene. Toxicol Lett 2014; 229: 101-10.

[13] Huang Y, Xia L, Wu Q, Zeng Z, Huang Z, Zhou S, Jin J, Huang H. Trichloroethylene Hypersensitivity Syndrome Is Potentially Mediated through Its Metabolite Chloral Hydrate. PLoS One 2015; 10: e0127101.

[14] Issaq HJ, Xiao Z, Veenstra TD. Serum and plasma proteomics. Chem Rev 2007; 107: 3601-20.

[15] Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. Nature 2008; 452: 571-9.

[16] Ray S, Reddy PJ, Jain R, Gollapalli K, Moiyadi A, Srivastava S. Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. Proteomics 2011; 11: 2139-61.

[17] Liang T, Wang N, Li W, Li A, Wang J, Cui J, Liu N, Li Y, Li L, Yang G, Du Z, Li D, He K, Wang G. Identification of complement C3f-desArg and its derivative for acute leukemia diagnosis and minimal residual disease assessment. Proteomics 2010; 10: 90-8.

[18] Villanueva J, Martorella AJ, Lawlor K, Philip J, Fleisher M, Robbins RJ, Tempst P. Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. Mol Cell Proteomics 2006; 5: 1840-52.

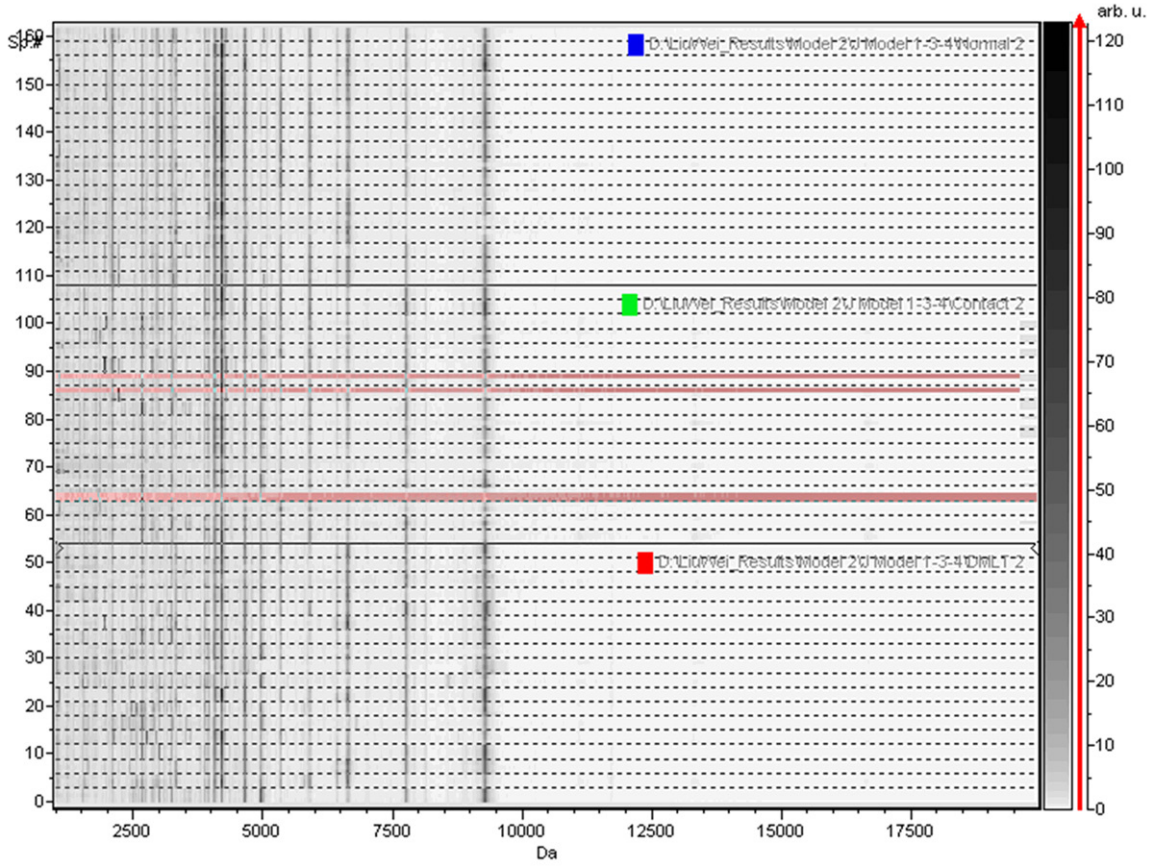
[19] Pusch W, Flocco MT, Leung SM, Thiele H, Kostrzewa M. Mass spectrometry-based clinical proteomics. Pharmacogenomics 2003; 4: 463-76.

[20] Goldman R, Ressom HW, Abdel-Hamid M, Goldman L, Wang A, Varghese RS, An Y, Loffredo CA, Drake SK, Eissa SA, Gouda I, Ezzat S, Moiseiwitsch FS. Candidate markers for the detection of hepatocellular carcinoma in low molecular weight fraction of serum. Carcinogenesis 2007; 28: 2149-53.

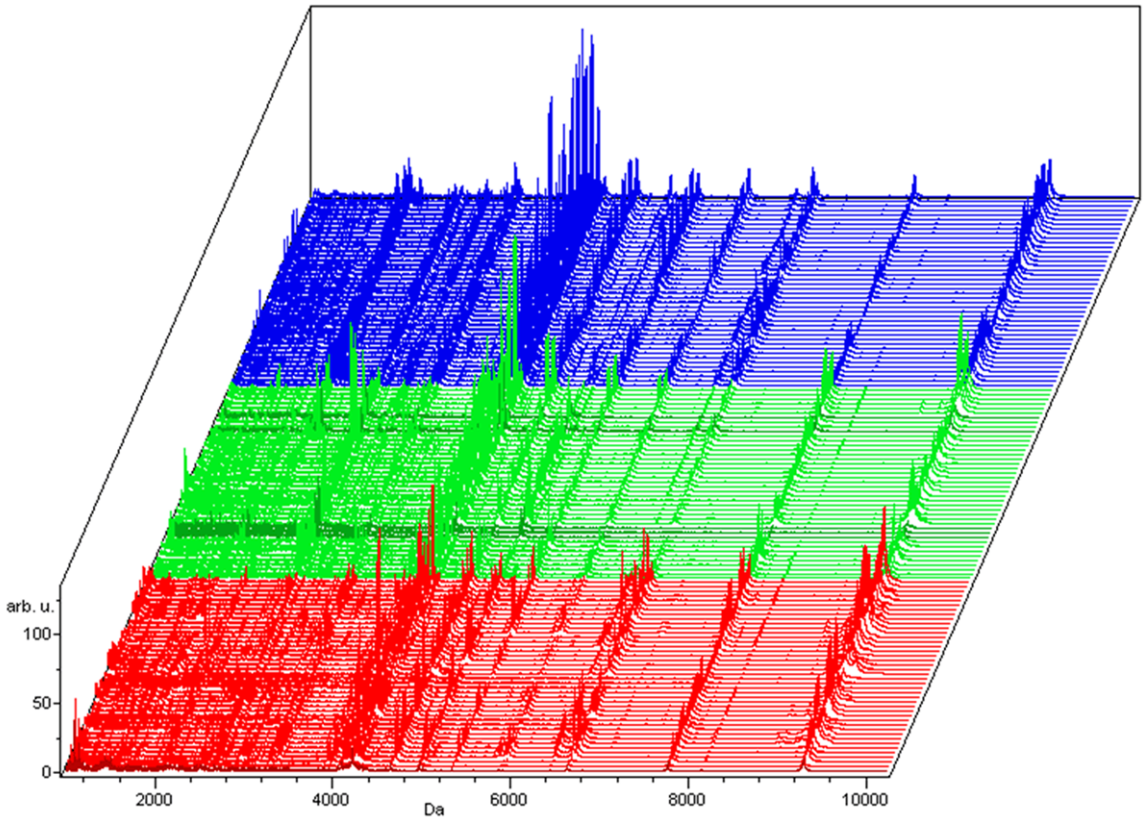
[21] Ketterlinus R, Hsieh SY, Teng SH, Lee H, Pusch W. Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. Biotechniques 2005; Suppl: 37-40.

[22] Singthong S, Pakkong P, Choosang K, Wongsanit S. Occupational health risks among trichloroethylene-exposed workers in a clock manufacturing factory. Glob J Health Sci 2014; 7: 161-72.

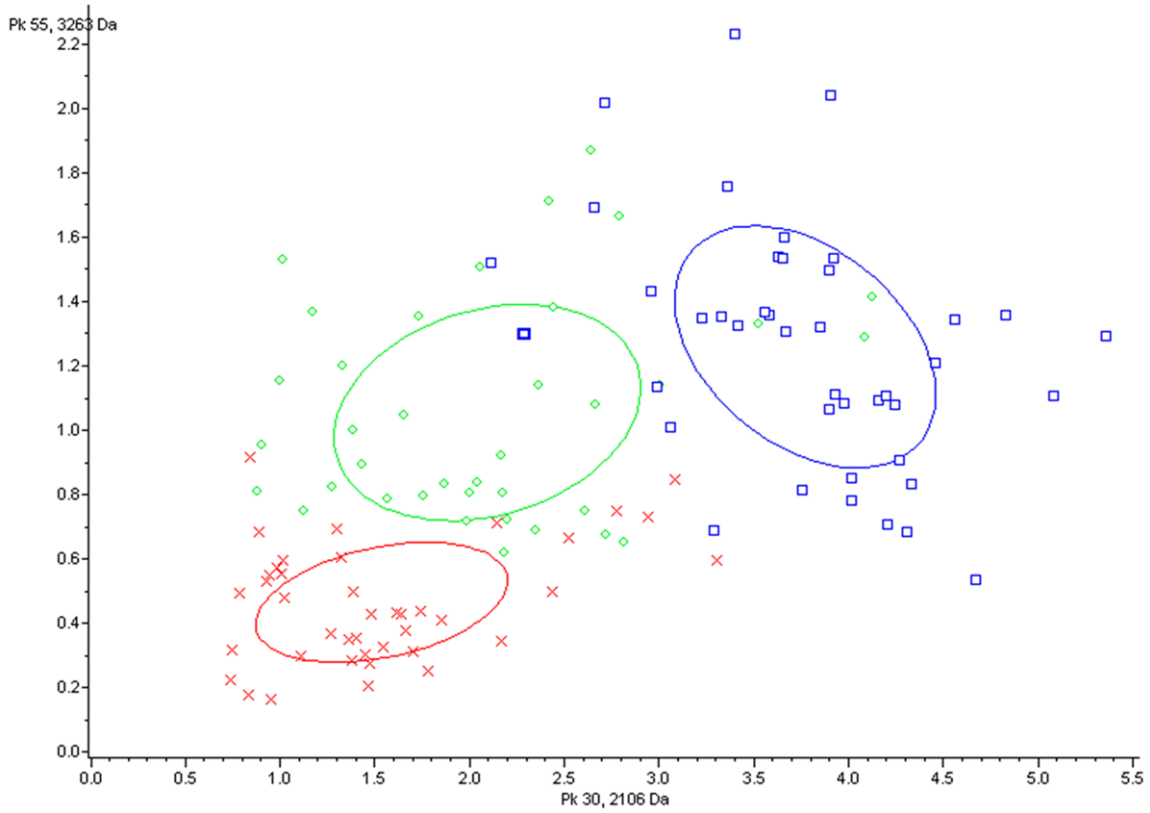
[23] Watanabe H. Hypersensitivity syndrome due to trichloroethylene exposure: a severe generalized skin reaction resembling drug-induced hypersensitivity syndrome. J Dermatol 2011; 38: 229-35.



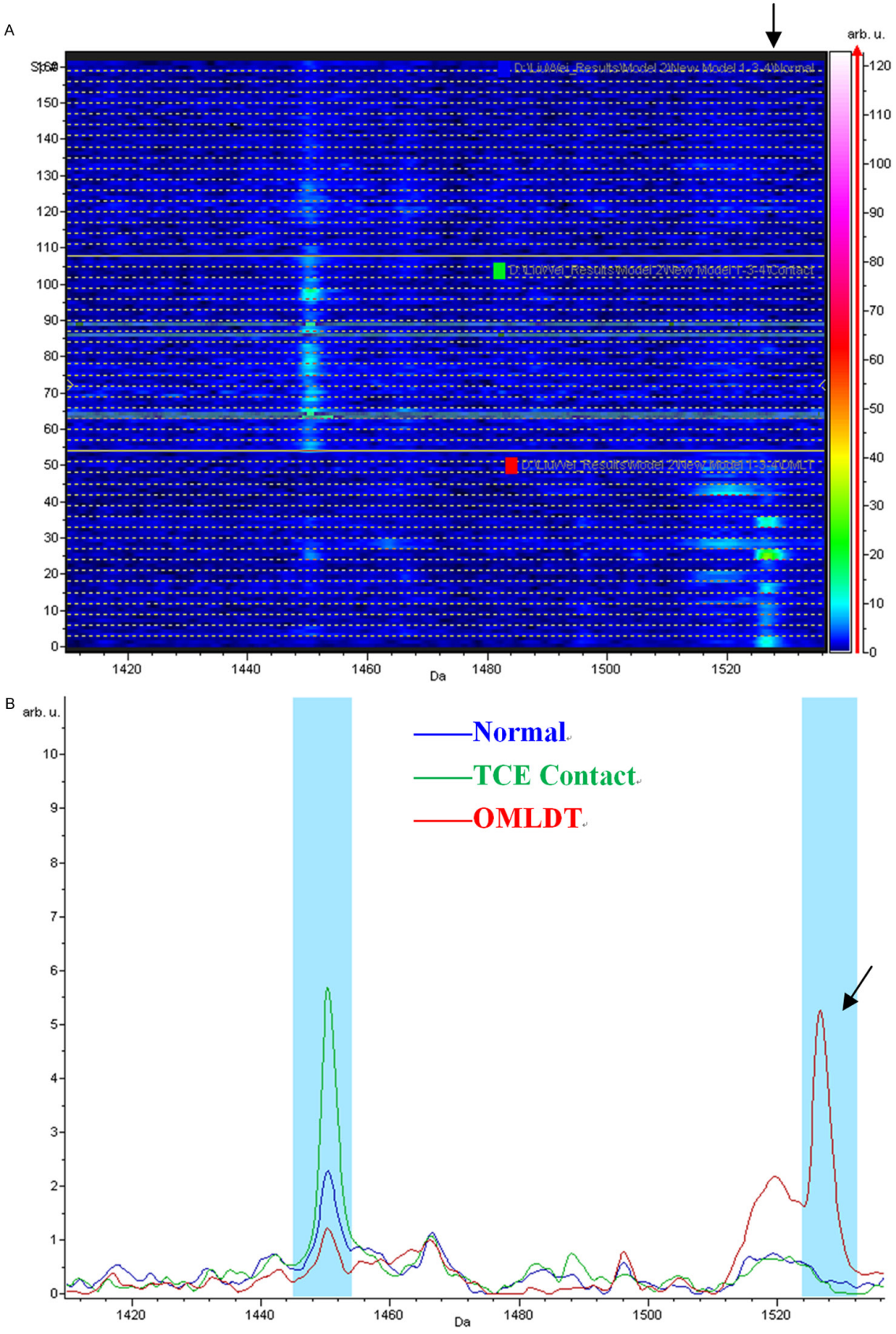
Supplementary Figure 1. Virtual gel view comparison of three groups. The above (blue) was normal group, middle (green) was TCE contact group and below (red) was OMLDT group. The x-axis records the m/z value. The left y-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a color code. The color bar and the right y-axis indicate the relation between the color a peak is displayed with and the peak intensity in arbitrary units.

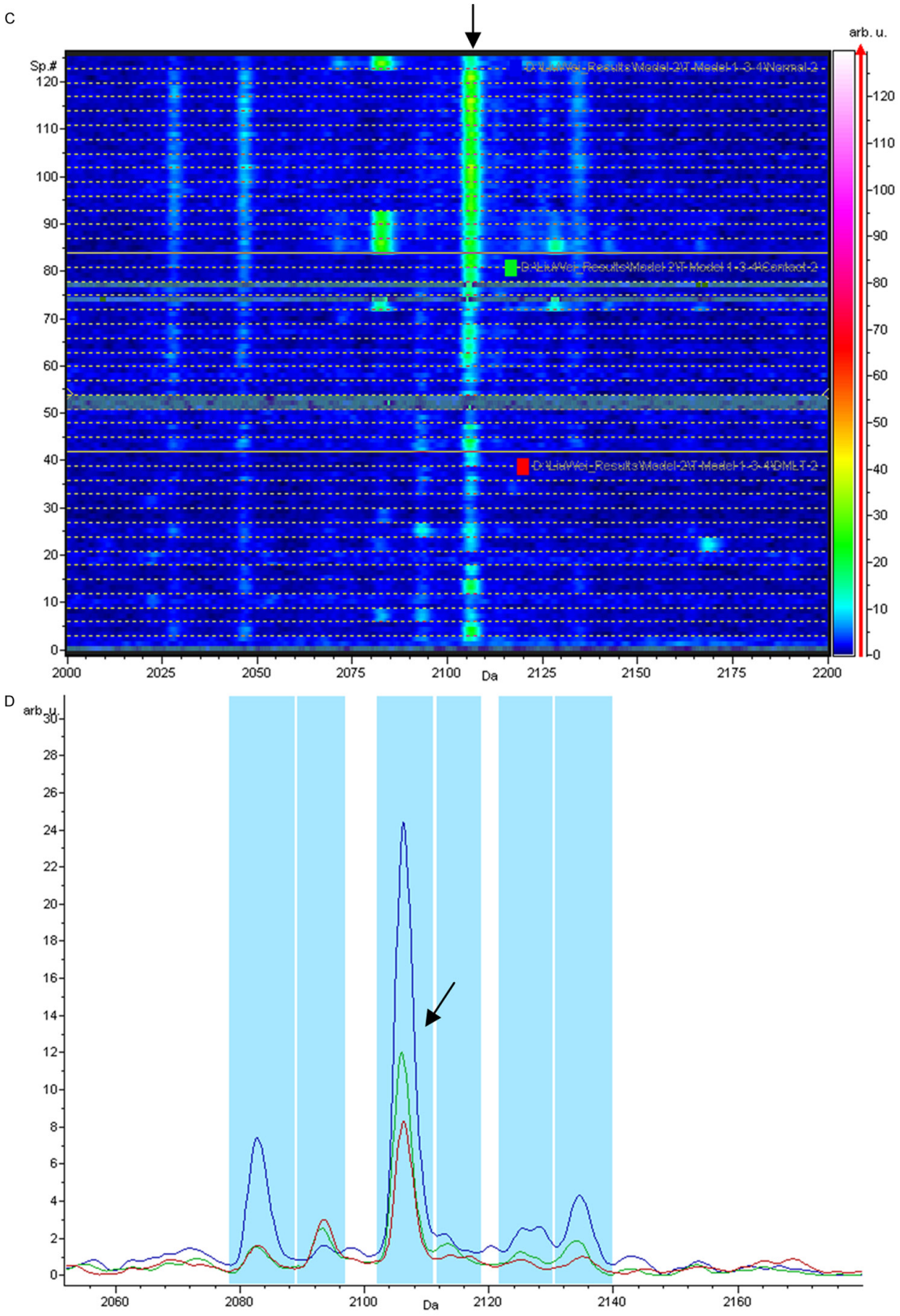


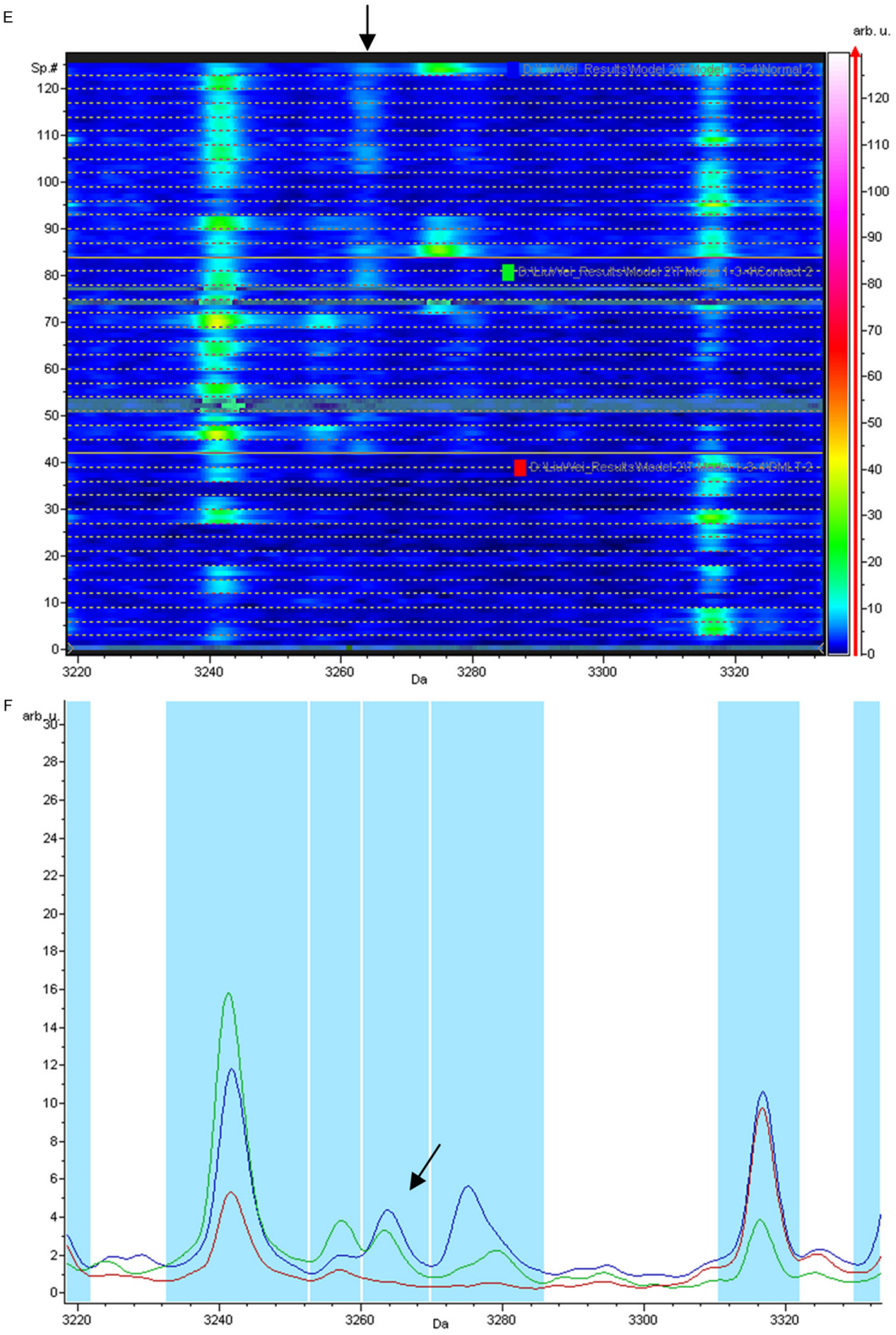
Supplementary Figure 2. Audio-visual comparison among OMLDT (red), TCE contact (green) and normal (blue) group in Stack View. The x-axis records the m/z value, the y-axis the peak intensity in arbitrary units and the z-axis the loading order.



Supplementary Figure 3. 2D peak distribution view of peptide with m/z 2106 (x axis) and 3263 (y axis) among OMLDT patients (Red Cross), TCE contact controls (Green Circle) and normal controls (Blue Square). The discrimination features of two selected peptides were generated by ClinProTools bioinformatics software. Value was represented peptide abundance ratio and showed significant difference among three groups. The ellipses represented the standard deviation of the class average of the peak areas/intensities.


A
arb. u.
Da
B
arb. u.
Normal
TCE Contact
OMLDT
Da


C
arb. u.
Sp.#
Da
D
arb. u.
Da



Supplementary Figure 4. Gel view and mass spectrum of m/z 1526 (A, B), 2106 (C, D) and 3263 (E, F) Da. In the left map (A, C, E), the x-axis records the m/z value. The left y-axis displays the running spectrum number. The peak intensity is expressed by a color code. The color bar and the right y-axis indicate the r elation between the color a peak is displayed with and the peak intensity in arbitrary units. In the right map (B, D, F), the x-axis records the m/z value and the y-axis the peak intensity in arbitrary units. The blue represents normal group, green was TCE contact group and red was OMLDT group. The arrows indicate the specific peaks.