UC San Diego
UC San Diego Previously Published Works

Title

Mass Spectrometry-Based Chemical Cartography of a Cardiac Parasitic Infection

Permalink

Journal

Analytical Chemistry, 89(19)

ISSN

0003-2700

Authors

McCall, Laura-Isobel
Morton, James T
Bernatchez, Jean A
et al.

Publication Date

2017-10-03

DOI

10.1021/acs.analchem.7b02423

Peer reviewed

Mass Spectrometry-Based Chemical Cartography of a Cardiac Parasitic Infection

Laura-Isobel McCall[1,*,#], **James T. Morton**[2], **Jean A. Bernatchez**[1], **Jair Lage de Siqueira-Neto**[1], **Rob Knight**[2,3,4], **Pieter C. Dorrestein**[1,4,5], and **James H. McKerrow**[1]

[1]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA 92093

[2]Department of Computer Science, University of California San Diego, La Jolla, CA 92093

[3]Department of Pediatrics, University of California San Diego, La Jolla, CA 92093

[4]Center for Microbiome Innovation, University of California San Diego, La Jolla, CA 92093

[5]Collaborative Mass Spectrometry Innovation Center, University of California San Diego, La Jolla, California 92093

[#]Present address: Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma, 73019-5251

Abstract

Trypanosoma cruzi parasites are the causative agents of Chagas disease, a leading infectious form of heart failure whose pathogenesis is still not fully characterized. In this work, we applied untargeted liquid chromatography-tandem mass spectrometry to heart sections from *T. cruzi*-infected and uninfected mice. We combined molecular networking and 3-dimensional modeling to generate chemical cartographical heart models. This approach revealed for the first time preferential parasite localization to the base of the heart and regiospecific distributions of nucleoside derivatives and eicosanoids, which we correlated to tissue-damaging immune responses. We further detected novel cardiac chemical signatures related to the severity and ultimate outcome of the infection. These signatures included differential representation of higher *vs* lower molecular weight carnitine and phosphatidylcholine family members in specific cardiac regions of mice infected with lethal or non-lethal *T. cruzi* strains and doses. Overall, this work provides new insights into Chagas disease pathogenesis and presents an analytical chemistry approach that can be broadly applied to the study of host-microbe interactions.

Graphical Abstract
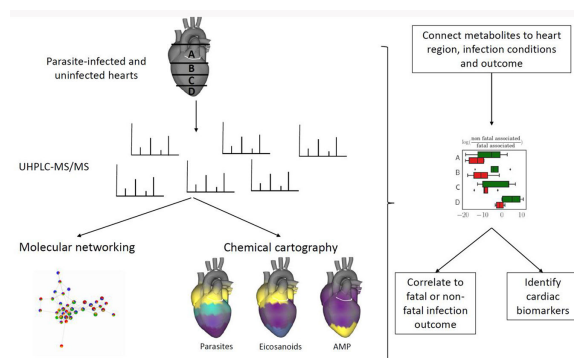
Chagas disease is a parasitic infection caused by *Trypanosoma cruzi*. Symptomatic Chagas disease usually presents with cardiac manifestations, leading in some cases to fatal heart failure. Infection with *T. cruzi* may also lead to digestive system pathology such as megaoesophagus and megacolon [1]. Chagas disease is endemic to Latin America. However, transmission has been reported in several states in the USA. Infected individuals are also now found worldwide due to emigration, including over 360,000 in the United States and Spain [2]. Worldwide prevalence of *T. cruzi* infection is approximately six to seven million people, 30-40% of which will develop symptomatic Chagas disease [3].

The pathogenesis of Chagas disease is still not fully understood. Until recently, Chagas disease was thought to be primarily autoimmune. However, more sensitive parasite detection methods have shown that persistent parasite presence is required. Disease is thought to result from a combination of elevated pro-inflammatory cytokine release, CD8+ T cell-mediated cytotoxicity and reduced regulatory immune mechanisms leading to tissue damage [4]. Parasite strain, dose, route of infection, host immune status and host genetic background can further modulate the outcome of infection, but the molecular details of their role in disease progression are still unknown [3,5–8]. Clinicians are still unable to predict which patients will develop severe disease and which will remain asymptomatic, so that current treatment guidelines recommend treating all seropositive patients aged 50 and under, even though the standard of care drug, benznidazole, causes significant adverse effects [9]. There is therefore an urgent need for a better understanding of Chagas disease pathogenesis, and for the identification of markers of infection outcome.

Fatal cases of Chagas disease often display a characteristic apical aneurysm at the bottom of the heart upon autopsy [10]. We therefore sought to investigate the spatial relationship between parasite and host, focusing on the small molecule interplay involved in parasite distribution and outcome of infection, by applying untargeted liquid chromatography-tandem mass spectrometry analysis to sequential transverse heart sections, in a mouse model of Chagas disease. Varying the parasite dose and parasite strain allowed us to predictably modulate the infection outcome, while molecular networking and principal balance analysis of this dataset enabled identification of affected metabolites. This chemical cartography approach enabled us to discover preferential parasite distribution to the top (base) of the heart, in association with intrinsic differential metabolite patterning across the heart. Furthermore, we provide the first evidence that small molecule profile can be used to differentiate between severe and

non-severe infections. Finally, we identified new pathways associated with infection outcome that could be targeted for anti-Chagasic therapy.

Experimental section

In vivo *experimentation*

Five-week old C3H/HeJ male mice (Jackson Laboratories) were infected intraperitoneally with culture-derived *T. cruzi* trypomastigotes of strains CL, X10/4 or X10/7, or left uninfected (see Supplemental experimental procedures for parasite culture methods). Each experimental group included at least 5 mice. Mice were euthanized 12, 90 and 145 days post-infection and perfused with phosphate-buffered saline to remove circulating trypomastigotes. Hearts were transversely sectioned and snap frozen in liquid nitrogen. All vertebrate animal studies were performed in accordance with the USDA Animal Welfare Act and the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the University of California San Diego Institutional Animal Care and Use Committee (protocol S14187). Samples collected from independent experimental infections with X10 strains and strain CL were analyzed jointly to identify strain-independent and mouse litter-independent commonalities in infection outcome. Additional controls included axenic *T. cruzi* epimastigotes (vector stage), extracellular culture-derived *T. cruzi* trypomastigotes (mammalian stage), intracellular *T. cruzi* amastigotes, uninfected host cells (mouse C2C12 myoblasts), and uninfected C2C12 culture supernatant.

Metabolite extraction

Metabolite extraction was performed by a two-step protocol as described by Want *et al*[11]. Briefly, heart sections were homogenized in water with steel beads using a Qiagen TissueLyzer at 25 Hz for 3 min. One-half of the homogenate was set aside for cytokine assays, DNA extraction and qPCR (see Supplemental experimental procedures). Methanol was added to the remainder, to a final concentration of 50%. Samples were further homogenized for 3 min at 25 Hz and centrifuged at 16,000g for 10 min. The centrifugation supernatant, subsequently referred to as the aqueous extract, was dried in a vacuum concentrator, and the dried pellet frozen at −80°C until LC-MS/MS analysis. The centrifugation pellet was resuspended in 3:1 dichloromethane:methanol, homogenized at 20 Hz for 5 min, and centrifuged at 16,000g for 10 min. This latter centrifugation supernatant, subsequently referred to as the organic extract, was air-dried at room temperature and frozen at −80°C until UHPLC-MS/MS analysis.

UHPLC-MS/MS analysis

Dried samples were resuspended in 50% methanol spiked with sulfadimethoxine as injection control. Liquid chromatography and mass spectrometry were performed as previously described [12]. Samples were analyzed in a single experimental run to prevent batch effects. Uninfected samples and samples from infection with different CL doses were randomized, as were uninfected samples and samples from different X10 strains. Samples were also randomized in terms of section position. Liquid chromatography separation was performed on an UltiMate 3000 UHPLC system (Thermo Scientific), using a 1.7 μm C8 (50 × 2.1 mm)

UHPLC column (Phenomenex). Chromatography was performed with water + 0.1% formic acid as mobile phase A and acetonitrile + 0.1% formic acid as mobile phase B, at a flow rate of 0.5 mL/min. Column temperature was 40 °C. LC conditions were determined empirically for aqueous and organic extracts. The LC gradient for the aqueous extract was: 0-1 min, 2% B; 1-1.5 min ramp up linearly to 40% B; 1.5-4 min ramp up linearly to 98% B; 4-5 min hold at 98% B; 5-6 min ramp down to 2% B; 6-7 min hold at 2% B. The LC gradient for the organic extract was: 0-1 min, 2% B; 1-1.5 min ramp up linearly to 60% B; 1.5-5.5 min ramp up linearly to 98% B; 5.5-7.5 min hold at 98% B; 7.5-8.5 min ramp down to 2% B; 8.5-10.5 min hold at 2% B. MS/MS analysis was performed on a Maxis Impact HD QTOF mass spectrometer (Bruker Daltonics). Ions were generated by electrospray ionization and MS spectra acquired in positive ion mode. Daily calibration was performed with ESI-L Low Concentration Tuning Mix (Agilent Technologies). Hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazene (Synquest Laboratories), m/z 922.009798, was present throughout the run and used as internal calibrant (lock mass).

Instrument parameters were as follows: nebulizer gas pressure, 2 Bar; Capillary voltage, 4,500 V; ion source temperature, 200°C; dry gas flow, 9.0 L/min; spectra rate acquisition, 3 spectra/s. MS/MS data was collected by fragmentation of the ten (organic extract) and seven (aqueous extract) most intense ions, in mass range 80-2,000 m/z, with active exclusion after 4 spectra and release after 30s. Mass ranges representing common contaminants and the lock masses were also excluded (exclusion list 123.59-124.59, 143.50-144.50, 159.47-160.47, 182.49-183.49, 216.61-217.61, 309.83-310.83, 337.50-338.50, 359.50-360.50, 622.00-622.05, 921.50-925.50). Ramped collision-induced dissociation energy parameters ranged from 10-50 eV.

LC-MS/MS data analysis

Raw data files were lock mass-corrected and converted to mzXML format using Compass Data analysis software (Bruker Daltonics). MS1 features were identified and normalized to the injection control (sulfadimethoxine) using an OpenMS-based [13] workflow (Optimus, July 21 2016 version, https://github.com/alexandrovteam/Optimus, see Table S-1 for parameters) and restricted to features with MS2 data available. Subsequent total ion current (TIC) normalization was performed in R studio. Principle coordinate analysis (PCoA) was performed on these MS1 features using the Bray-Curtis-Faith dissimilarity metric in our in-house tool ClusterApp and visualized in EMPeror[14]. Three-dimensional modeling was performed with 'ili (http://ili.embl.de/), using a three-dimensional heart model from 3DCADBrowser.com (http://www.3dcadbrowser.com/).

Molecular networking was performed using the Global Natural Products Social Molecular Networking platform (GNPS)[15], with the following parameters for experimental spectra and library searches: parent mass tolerance 0.02 Da, MS/MS fragment ion tolerance 0.02 Da, cosine score > 0.6, > 4 matched peaks. Analog search against the library allowed a maximum mass shift of 200 Da. Molecular networks were visualized with Cytoscape 3.4.0. Venn diagrams were generated from Cytoscape data at http://bioinformatics.psb.ugent.be/webtools/Venn/. Additional structural predictions were performed using CSLFingerID [16] and the LIPID MAPS resources [17].

Random forest analysis [18] was performed in R studio. Validation of the random forest classifier for infection outcome was performed by splitting the dataset into training and test datasets, building the classifier on the training data, and then evaluating its performance on the test dataset. All boxplots represent first quartile, median and third quartile, with whiskers extending no more than 1.5 times the interquartile range. Notches extend to +/−1.58 interquartile range/sqrt(n). Non-overlapping notches indicate significantly different medians (95% confidence). Receiver operating characteristic (ROC) curves were generated in GraphPad Prism 6.

Principal balance analysis was performed using Gneiss [19,20]. Partitions of metabolites were defined using Ward's Hierarchical Clustering with a proportionality distance metric. This clustering approximately groups together metabolites that are correlated with each other. With these partitions, the ilr transform applied to obtain principal balances. A multivariate response linear model accounting for the sample dependence with respect to the sampling site was performed directly on the balances using the following equation, testing for the effects of parasite burden, parasite dosage, sampling site along the heart and outcome.

$$y = \beta_0 + \beta_1 x_{burden} + \beta_2 x_{position} + \beta_3 x_{dose} + \beta_4 x_{outcome}$$

where y is the metabolite composition, $\beta_0, \ldots, \beta_4$ are the estimated regression coefficients, X_{burden} is the measured endpoint cardiac parasite burden, $x_{position}$ is the section position in the heart, x_{dose} is the amount of parasite injected at time 0 and $x_{outcome}$ is the mouse survival outcome.

Data availability

Metabolomics data has been deposited in MassIVE (http://massive.ucsd.edu/, accession #MSV000080450). Molecular networks can be accessed at http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=ed620c942bbf4c15a3a1d418150de62c (specific uninfected 12 day sample (mouse 4 section A, aqueous extract, no window filtering), http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c9ad00413f75402e9d56f4d93e8e5142 (uninfected 12 day samples, aqueous extract), http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=b900edba13cb4e649280605d58467e6b (all samples, aqueous extract), and http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=b92349791ebf4571b792d54be8edd55d (all samples, organic extract). Code for principal balance analysis of this dataset has been deposited at https://github.com/knightlab-analyses/chagas-heart-metabolomics.

Results and discussion

Spatial mapping of cardiac infection by T. cruzi

Chagas disease presents with characteristic cardiac lesions such as apical aneurysms at the bottom of the heart[10], but the mechanisms leading to this targeted damage are currently unknown. To investigate the cause of these characteristic lesions, we performed metabolomic and parasitological analysis of sequential transverse heart sections in uninfected mice and in mice infected with *T. cruzi* strains CL, X10/4 or X10/7. Parasite burden was predominantly in the base ("top") of the heart under fatal infection conditions

(Fig. 1a, b), whereas low dose infection with strain CL (acute stage) presented with a more balanced parasite distribution. Infection with strain X10/4 (non-fatal) showed a variable parasite distribution in the acute stage, with a more uniform parasite distribution across the heart in the chronic stage. These observations suggest that parasite localization away from the heart base could be associated with mouse survival.

The pro-inflammatory cytokine IFNγ plays a key role in parasite killing, at the cost of tissue damage in Chagas disease [8]. We therefore investigated IFNγ distribution across the heart. Total and apical IFNγ levels were comparable 12 days post-infection in mice infected with strains X10/4, X10/7 and high dose strain CL, while IFNγ levels were significantly lower in the other cardiac sections, indicating a predominant contribution from the apex of the heart to total cardiac IFNγ under those infection conditions (2-way ANOVA, Fig. 1c, d). In those groups, IFNγ levels and parasite burden showed opposite distributions. In contrast, IFNγ levels were equally distributed between sections in mice infected with 1,000 parasites of strain CL, even though parasites were still predominantly localized to the heart base, supporting I FNy-independent factors regulating parasite distribution.

Given this targeted parasite distribution (IFNγ-independent for mid-dose strain CL), we hypothesized that there were inherent chemical differences between heart regions that could explain parasite tropism within the heart. These could include specific nutrients required by the parasite, immune mediators, or other restriction factors. Metabolites represent the final output of mRNA and protein expression as well as protein activity, so that metabolomic studies can provide information that is very closely related to observed phenotype [21]. We therefore extracted metabolites from sequential transverse heart sections via a two-step aqueous and organic solvent extraction. This approach provides a broad representation of polar and non-polar metabolites [11]. Extracted metabolites were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS), and multivariate statistical analyses were performed. Overall distinct metabolite signatures were identified for the base and apex of the heart, irrespective of infection outcome or collection timepoint (Fig. 1e, Fig S-1). These results suggest that the cardiac environment encountered by the parasites differs across heart regions.

To deconvolute infection-associated changes in cardiac metabolites from inherent cardiac regiospecific characteristics, we focused on uninfected heart samples. Signature features differing between cardiac regions in uninfected mice were determined by random forest classification [18]. To obtain putative structural information for these results, we performed molecular networking using the Global Natural Products Social Molecular Networking platform to computationally align and group experimental and reference MS/MS spectra based on pattern similarity [15]. Overall, 31,383 spectra were used to generate a network of 963 merged consensus spectra (nodes). 205 of these nodes did not show spectral similarity to other nodes (self-looped), while the remainder of the spectra were grouped into 66 clusters of related spectra, indicating shared chemical groups (Fig. S-2c).

Eight of the top 30 differential chemical features identified by random forest on the aqueous extract in uninfected mice are eicosanoid family members and precursors (level three annotation according to the 2007 metabolomics standards initiative [22]; Table 1, Fig. S-3), an

over-representation compared to their frequency in the full dataset. We therefore compared total feature abundances of eicosanoid derivatives and precursors identified in this dataset, and observed elevated levels at the base of the heart compared to the apex (Fig. 1f, g). Eicosanoids produced by lymphocytes, macrophages and parasites have been implicated in Chagas disease pathogenesis [23]. Prostaglandins promote *T. cruzi* host cell invasion [24] and intracellular replication [25], so the higher eicosanoid levels at the base of the heart may partially account for the higher parasite burden observed in that region (compare Fig. 1f to Fig. 1b). Prostaglandins in the context of Chagas disease have also been shown to inhibit the production of antiparasitic TNFα and IFNγ [26,27], although none of the top 30 differential features identified by random forest were correlated with IFNγ levels (Fig. S-4).

An additional three differential features in the aqueous extract from uninfected mice are annotated as adenine nucleosides and derivatives, *m/z* 362.087 RT 44 s was elevated at the base of the heart, while the remaining two features were elevated at the apex of the heart (Fig. 1j, k, Fig. S-5a). *m/z* 362.087 is annotated as AMP with an extra methyl group, and *m/z* 348.071 as AMP. m/z 464.082 is a putative adenylosuccinic acid, an intermediate in the conversion of inosine monophosphate to AMP [28]. *m/z* 364.066, also identified by random forest, is putatively annotated as GMP. *T. cruzi* is a purine auxotroph, dependent on host purines for metabolic functions, but *T. cruzi* imports purine nucleosides rather than nucleobases or nucleotides [29]. We therefore investigated the distribution of adenosine across the heart. Similar to AMP, adenosine levels were elevated at the apex of the heart (Fig. 1h, i). However, transcriptome analysis of intracellular amastigotes indicates preferential purine uptake and salvage via the guanosine/guanine branch of the purine salvage pathway [30], rather than the adenosine/adenine branch. We therefore assessed guanosine levels and found them comparable across the heart (Fig. 1h, i). Therefore, the limited adenosine levels at the base of the heart are likely compensated by the comparable guanosine distribution across the entire heart, and the differential adenine nucleoside levels across the heart are therefore unlikely to affect *T. cruzi* growth directly. In contrast, host nucleoside monophosphates also play signaling roles. AMP in particular regulates the activity of AMP-activated protein kinase (AMPK), and inhibition of AMPK promotes parasite growth [29]. Increased AMP levels at the apex of the heart would therefore lead to increased AMPK activity and parasite growth restriction. In contrast, structure-based molecular modeling predicts an inability of methylated AMP to bind the AMP binding site on AMPK (Fig. S-5f), so that methylated AMP would not activate AMPK, allowing for unrestricted parasite growth.

Distinct regiospecific chemical signatures associated with fatal and non-fatal infection outcomes

Although apical aneurysms are pathognomonic for Chagas disease, human infections with *T. cruzi* are not uniformly fatal [10]. Parasite factors, including dose, route of infection, and parasite strain, as well as host genotype and immune status all influence the outcome of infection with *T. cruzi* [3,5–8]. In susceptible mouse models, high dose infection with strain CL is fatal within a few weeks, while ~50% of C3H/HeJ mice infected with 1,000 trypomastigotes survive the acute stage and develop cardiac pathology. Infection with lower dose is associated with mouse survival [31]. Likewise, in C3H/HeJ male mice, strain X10/7 is rapidly lethal, while mice infected with strain X10/4 survive into the chronic stage [32] (Fig.

S-6). Given that cardiac pathology develops early after infection [8,33], we investigated whether we could identify early chemical signatures that were associated with fatal *vs* non-fatal outcomes 12 days post-infection.

Irrespective of infecting strain, outcome was strongly associated with cardiac parasite burden at time of sacrifice (Fig. 2c). Distinct clustering by infection outcome was observed only for the organic extract (Fig. 2a, p<0.001, Adonis; Fig. S-7), so all subsequent analysis of outcome-associated metabolites focused on the organic extract exclusively. We initially performed random forest analysis on all cardiac sections to determine whether we could use our metabolomics data to classify samples based on infection outcome. Our classifier correctly classified the majority of samples from 100% fatal and non-fatal infection conditions; in particular, no non-fatal samples were misidentified as fatal samples (Fig. 2b, Table S-2). Identification of differential features was performed as above by molecular networking on the full organic extraction dataset (infected and uninfected). Overall, 437,259 spectra were used to generate a network of 3,608 merged consensus spectra (nodes), including 245 clusters of related spectra with shared chemical groups (Fig. S-2a). Several of these key differential factors for this classification are members of the phosphatidylcholine or sphingolipid families, including *m/z* 606.619 RT 448s (Fig. 2d), even though total levels of molecules identified as phosphatidylcholines were comparable between outcomes (Fig. S-8). *m/z* 606.619 RT 448s was computationally identified as the sphingolipid N-(docosanoyl)-1-deoxysphing-4-enine. Additional factors differentiating between fatal and non-fatal outcomes in our random forest classifier include m/z 512.468 RT 276s, which is found in a network of nodes related to acetyl-carnitine and long-chain acylcarnitines (Fig. 2e, Fig. S-9). ROC curve analysis indicates that each of these metabolites can be used to differentiate between fatal and non-fatal infections (p<0.01 and p<0.005, respectively).

Outcome of infection is likely due to a combination of factors, rather than individual metabolites. To address this issue, analyses based on principal balances were performed on all cardiac sections [19]. This technique enables the application of standard statistical analysis directly to partitions of molecules instead of individual molecules, allowing for investigation of interactions of large groups of molecules. A multivariate response linear model was applied directly on the balances, testing for the effects of endpoint cardiac parasite burden, parasite dose at the time of infection, heart section position and outcome. Based on these analyses, it was found that two groups of metabolites characterized a clear separation between fatal and non-fatal samples, especially for samples from the base of the heart for low dose strain CL *vs* high dose strain CL and the apex of the heart for non-fatal strain X10/4 *vs* fatal strain X10/7 (Fig. 2f, g, Tables S-3, S-4 and S-5). These differences in position between X10/4 and CL infection could be due to parasite persistence at the base of the heart (section A) with strain CL during the transition from acute to chronic stage, whereas parasites shift away from the base of the heart to central cardiac sections during the transition from acute to chronic stage in strain X10/4 infection. This suggests a stronger role for non-basal sections of the heart during chronic infection with strain X10/4 than with strain CL, which is mirrored by our balance tree analysis.

Differential molecules identified in our balance tree analysis include phosphatidylcholine, eicosanoid and carnitine family members. A few of these differential molecules are parasite-

derived (Fig. S-10, Table S-5), but the vast majority are host-derived. Long- and short-chain acylcarnitines in particular segregated between fatal and non-fatal outcomes (Fig. 2h). Likewise, balance tree analysis revealed a shift in phosphatidylcholine sizes between outcomes (Fig. 2i), with an increased representation of higher molecular weight lipids in non-fatal outcomes. Long and short-chain acylcarnitines and fatty acids play differential roles in cardiac muscle energy metabolism, stress response, and oxidative damage management [34]. Phosphatidylcholines are major membrane components, but also play important roles in immune regulation [35,36]. Furthermore, modulations of serum phosphatidylcholine levels have been detected in coronary heart disease and non-infectious heart failure [37,38]. Parasite-derived phosphatidylcholines also induce platelet aggregation, a mechanism that is involved in Chagas disease pathogenesis [39].

Conclusions

The factors leading to the typical apical aneurysm in Chagas disease are currently unknown. In this work, we observed regiospecific chemical differences in heart tissue that could be correlated with parasite distribution and infection outcome. These factors include the distribution of specific eicosanoids, and nucleosides, as well as shifts in the relative abundance of long-chain *vs* short-chain acylcarnitine and phosphatidylcholine family members. Changes in phosphatidylcholines and carnitine family members have been observed in targeted studies of *T. cruzi*-infected animals compared to uninfected mice [40], in myocardial infarct [41,42], and in coronary heart disease [38,43,44], but this is the first time they have been tied to *T. cruzi* infection outcome. To the best of our knowledge, this is the first study to identify regiospecific small molecule differences across the heart that could be tied to disease, with potential impact on the pathology of other cardiovascular diseases. These discoveries were directly enabled by our chemical cartography approach. Such techniques could readily be applied to the spatial study of other host-microbe interactions, to uncover new pathways of microbial pathogenesis.

It is important to note that most of the detected molecules are un-annotated. Our interpretation of this LC-MS/MS dataset is therefore limited to characterized molecules, even though many other currently unidentified molecules and low-abundance molecules lacking MS2 data will probably influence the outcome of infection. Given the relative abundance of host cells compared to parasites (especially in the chronic stage of disease), the majority of the observed signals are host-derived, although several shared signatures were observed between heart sections and cultured host-free parasite stages (epimastigotes), while being absent from cultured mouse cardiac cells. These chemical signatures include a number of phosphatidylcholine family members. We also identified signatures from trypomastigotes and from amastigotes-containing host cells that are absent in uninfected mouse cells. These are either parasite-derived or infection-induced (Fig. S-10, Table S-5), and will likely also play a role in influencing the outcome of infection.

This work represents the first large-scale untargeted metabolomic study of Chagas disease progression, with unprecedented chemical coverage and spatial resolution. Predicting patient prognosis is currently impossible for Chagas disease, and our results represent the first successful usage of metabolite profile to classify samples based on the outcome of infection.

Future work will investigate whether we can expand these observations to plasma samples for patient diagnosis and staging. Finally, we identified several new pathways involved in parasite tropism and pathogenesis. These represent valuable avenues to be investigated for antiparasitic drug development. Overall, our results represent a novel combination of analytical chemistry tools with cutting-edge mass spectrometry data analysis techniques, to provide significant insight into this neglected parasitic infection.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

(1). McCall LI; McKerrow JH Trends Parasitol. 2014, 30, 342–349. [PubMed: 24946952]

(2). Gascon J; Bern C; Pinazo MJ Acta Trop. 2010, 115, 22–27. [PubMed: 19646412]

(3). Rassi A, Jr.; Rassi A; Marin-Neto JA Lancet 2010, 375, 1388–1402. [PubMed: 20399979]

(4). Rassi A, Jr.; Marin JAN; Rassi A Mem. Inst. Oswaldo Cruz 2017, 112, 224–235. [PubMed: 28225900]

(5). Borges DC; Araujo NM; Cardoso CR; Lazo Chica JE Immunology 2013, 138, 145–156. [PubMed: 23113506]

(6). Dias GB; Gruendling AP; Araujo SM; Gomes ML; Toledo MJ Exp. Parasitol.2013, 135, 511–517. [PubMed: 23994765]

(7). Frade AF; Pissetti CW; Ianni BM; Saba B; Lin-Wang HT; Nogueira LG; de Melo Borges A; Buck P; Dias F; Baron M; Ferreira LR; Schmidt A; Marin-Neto JA; Hirata M; Sampaio M; Fragata A; Pereira AC; Donadi E; Kalil J; Rodrigues V, et al. BMC Infect. Dis.2013, 13, 587. [PubMed: 24330528]

(8). Andrade DV; Gollob KJ; Dutra WO PLoS Negl. Trop. Dis. 2014, 8, e3010. [PubMed: 25077613]

(9). Bern C; Montgomery SP; Herwaldt BL; Rassi A, Jr.; Marin-Neto JA; Dantas RO; Maguire JH; Acquatella H; Morillo C; Kirchhoff LV; Gilman RH; Reyes PA; Salvatella R; Moore AC JAMA 2007, 298, 2171–2181. [PubMed: 18000201]

(10). Marin-Neto JA; Cunha-Neto E; Maciel BC; Simoes MV Circulation 2007, 115, 1109–1123. [PubMed: 17339569]

(11). Want EJ; Masson P; Michopoulos F; Wilson ID; Theodoridis G; Plumb RS; Shockcor J; Loftus N; Holmes E; Nicholson JK Nat. Protoc. 2013, 8, 17–32. [PubMed: 23222455]

(12). Bouslimani A; Porto C; Rath CM; Wang M; Guo Y; Gonzalez A; Berg-Lyon D; Ackermann G; Moeller Christensen GJ; Nakatsuji T; Zhang L; Borkowski AW; Meehan MJ; Dorrestein K; Gallo RL; Bandeira N; Knight R; Alexandrov T; Dorrestein PC Proc. Natl. Acad. Sci. U. S. A. 2015, 112, E2120–2129. [PubMed: 25825778]

(13). Sturm M; Bertsch A; Gropl C; Hildebrandt A; Hussong R; Lange E; Pfeifer N; Schulz-Trieglaff O; Zerck A; Reinert K; Kohlbacher O BMC Bioinf. 2008, 9, 163.

(14). Vazquez-Baeza Y; Pirrung M; Gonzalez A; Knight R Gigascience 2013, 2, 16. [PubMed: 24280061]

(15). Wang M; Carver JJ; Phelan VV; Sanchez LM; Garg N; Peng Y; Nguyen DD; Watrous J; Kapono CA; Luzzatto-Knaan T; Porto C; Bouslimani A; Melnik AV; Meehan MJ; Liu WT; Crusemann

M; Boudreau PD; Esquenazi E; Sandoval-Calderon M; Kersten RD, et al. Nat. Biotechnol. 2016, 34, 828–837. [PubMed: 27504778]

(16). Duhrkop K; Shen H; Meusel M; Rousu J; Bocker S Proc. Natl. Acad. Sci. U. S. A. 2015, 112, 12580–12585. [PubMed: 26392543]

(17). Fahy E; Sud M; Cotter D; Subramaniam S Nucleic Acids Res. 2007, 35, W606–612. [PubMed: 17584797]

(18). Breiman L Machine Learning 2001, 45, 5–32.

(19). Morton JT; Sanders J; Quinn RA; McDonald D; Gonzalez A; Vazquez-Baeza Y; Navas-Molina JA; Song SJ; Metcalf JL; Hyde ER; Lladser M; Dorrestein PC; Knight R mSystems 2017, 2.

(20). Pawlowsky-Glahn V; Egozcue JJ; Tolosana-Delgado R In Proceedings of the 4th International Workshop on Compositional Data Analysis, 2011.

(21). Breitling R; Bakker B; Barrett M; Decuypere S; Dujardin J-C In Genetics Meets Metabolomics, Suhre K, Ed.; Springer New York, 2012, pp 73–84.

(22). Sumner LW; Amberg A; Barrett D; Beale MH; Beger R; Daykin CA; Fan TW; Fiehn O; Goodacre R; Griffin JL; Hankemeier T; Hardy N; Harnly J; Higashi R; Kopka J; Lane AN; Lindon JC; Marriott P; Nicholls AW; Reily MD, et al. Metabolomics 2007, 3, 211–221. [PubMed: 24039616]

(23). Machado FS; Mukherjee S; Weiss LM; Tanowitz HB; Ashton AW Adv. Parasitol. 2011, 76, 1–31. [PubMed: 21884885]

(24). Malvezi AD; da Silva RV; Panis C; Yamauchi LM; Lovo-Martins MI; Zanluqui NG; Tatakihara VL; Rizzo LV; Verri WA, Jr.; Martins-Pinge MC; Yamada-Ogatta SF; Pinge-Filho P Mediators Inflamm. 2014, 2014, 580919. [PubMed: 25045211]

(25). Freire-de-Lima CG; Nascimento DO; Soares MB; Bozza PT; Castro-Faria-Neto HC; de Mello FG; DosReis GA; Lopes MF Nature 2000, 403, 199–203. [PubMed: 10646605]

(26). Borges MM; Kloetzel JK; Andrade HF, Jr.; Tadokoro CE; Pinge-Filho P; Abrahamsohn I Immunol. Lett. 1998, 63, 1–8. [PubMed: 9719432]

(27). Michelin MA; Silva JS; Cunha FQ Exp. Parasitol. 2005, 111, 71–79. [PubMed: 16009364]

(28). Marie S; van den Berghe G; Vincent M-F In Inborn Metabolic Diseases; Springer, 2016, pp 495–513.

(29). Caradonna KL; Engel JC; Jacobi D; Lee CH; Burleigh BA Cell Host Microbe 2013, 13, 108–117. [PubMed: 23332160]

(30). Li Y; Shah-Simpson S; Okrah K; Belew AT; Choi J; Caradonna KL; Padmanabhan P; Ndegwa DM; Temanni MR; Corrada Bravo H; El-Sayed NM; Burleigh BA PLoS Pathog. 2016, 12, e1005511. [PubMed: 27046031]

(31). Lewis MD; Francisco AF; Taylor MC; Jayawardhana S; Kelly JM Cell. Microb. 2016.

(32). Postan M; Dvorak JA; McDaniel JP Am. J. Trop. Med. Hyg. 1983, 32, 497–506. [PubMed: 6407346]

(33). Calvet CM; Meuser M; Almeida D; Meirelles MN; Pereira MC Exp.Parasitol. 2004, 107, 20–30. [PubMed: 15208034]

(34). Chintapalli SV; Jayanthi S; Mallipeddi PL; Gundampati R; Suresh Kumar TK; van Rossum DB; Anishkin A; Adams SH J. Biol. Chem. 2016, 291, 25133–25143. [PubMed: 27758871]

(35). Treede I; Braun A; Sparla R; Kuhnel M; Giese T; Turner JR; Anes E; Kulaksiz H; Fullekrug J; Stremmel W; Griffiths G; Ehehalt RJ Biol. Chem. 2007, 282, 27155–27164.

(36). Wu W; Shi X; Xu C Nat. Rev. Immunol. 2016, 16, 690–701. [PubMed: 27721483]

(37). Cheng M-L; Wang C-H; Shiao M-S; Liu M-H; Huang Y-Y; Huang C-Y; Mao C-T; Lin JF; Ho H-Y; Yang N-IJ Am. Coll. Cardiol. 2015, 65, 1509–1520.

(38). Ganna A; Salihovic S; Sundstrom J; Broeckling CD; Hedman AK; Magnusson PK; Pedersen NL; Larsson A; Siegbahn A; Zilmer M; Prenni J; Arnlov J; Lind L; Fall T; Ingelsson E PLoS Genet. 2014, 10, e1004801. [PubMed: 25502724]

(39). Gazos-Lopes F; Oliveira MM; Hoelz LV; Vieira DP; Marques AF; Nakayasu ES; Gomes MT; Salloum NG; Pascutti PG; Souto-Padron T; Monteiro RQ; Lopes AH; Almeida IC PLoS Negl. Trop. Dis. 2014, 8, e3077. [PubMed: 25101628]

(40). Girones N; Carbajosa S; Guerrero NA; Poveda C; Chillon-Marinas C; Fresno M PLoS Negl. Trop. Dis. 2014, 8, e3337. [PubMed: 25412247]

(41). Sansbury BE; DeMartino AM; Xie Z; Brooks AC; Brainard RE; Watson LJ; DeFilippis AP; Cummins TD; Harbeson MA; Brittian KR; Prabhu SD; Bhatnagar A; Jones SP; Hill BG Circ. Heart Fail. 2014, 7, 634–642. [PubMed: 24762972]

(42). Nam M; Jung Y; Ryu DH; Hwang GS Int. J. Cardiol. 2016.

(43). Bhattacharya S; Granger CB; Craig D; Haynes C; Bain J; Stevens RD; Hauser ER; Newgard CB; Kraus WE; Newby LK; Shah SH Atherosclerosis 2014, 232, 191–196. [PubMed: 24401236]

(44). Feng Q; Liu Z; Zhong S; Li R; Xia H; Jie Z; Wen B; Chen X; Yan W; Fan Y; Guo Z; Meng N; Chen J; Yu X; Zhang Z; Kristiansen K; Wang J; Xu X; He K; Li G Sci. Rep. 2016, 6, 22525. [PubMed: 26932197]
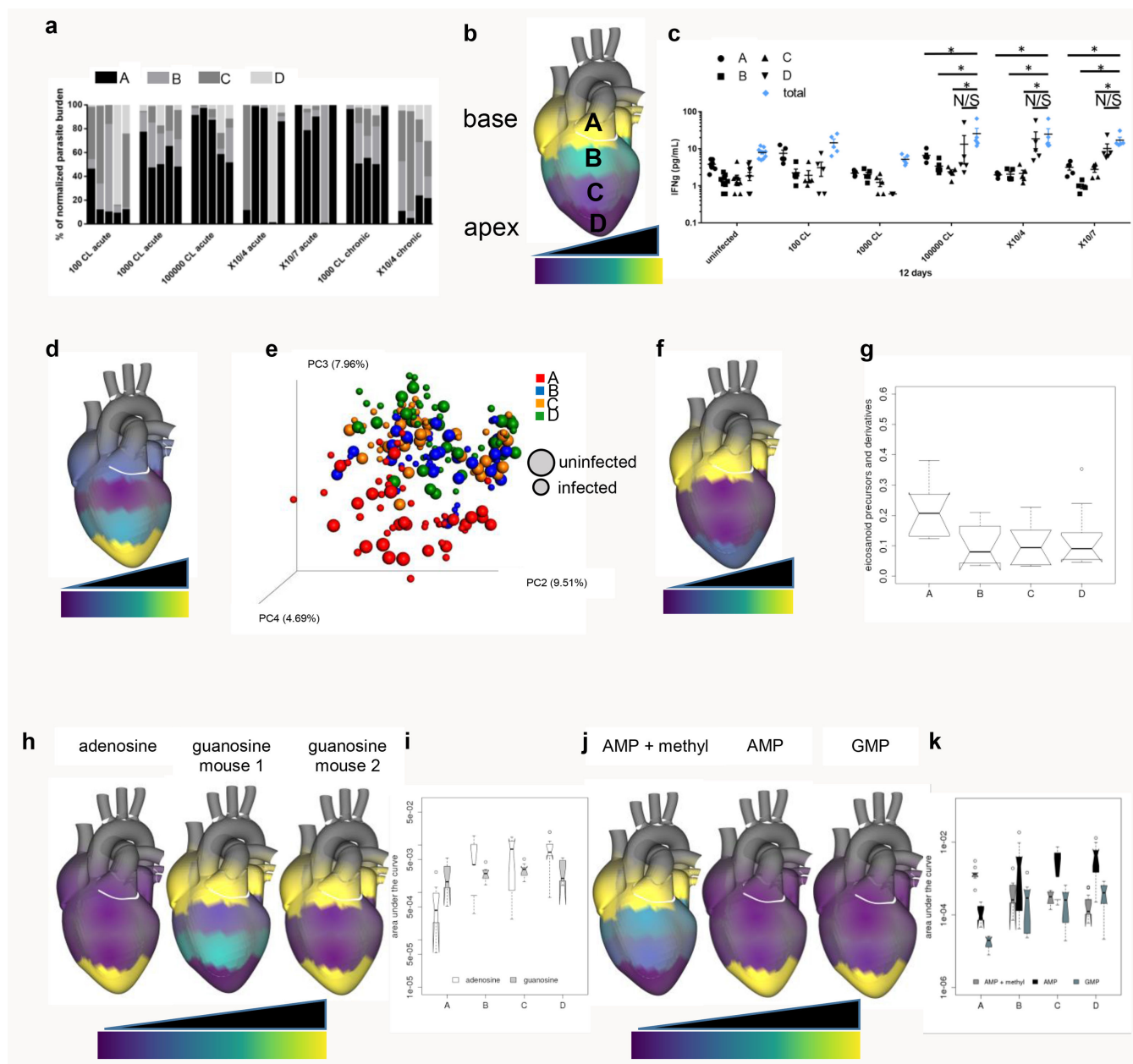
Figure 1. Preferential parasite location to the base of the heart is associated with distinct small molecule signatures.

Transverse heart sections were collected 12, 90 and 147 days post-infection from mice infected with *T. cruzi* strains CL, X10/4 and X10/7, and from matched uninfected controls. Parasite burden was determined by qPCR on extracted DNA. Metabolites were extracted with 50% methanol and analyzed by LC-MS/MS. Parasite burden was highest at the base of the heart (acute stage and chronic CL infection) overall (**a,** each bar represents a given mouse, grouped by infection condition) and in a representative mouse (**b**) Section positions (A to D, base to apex) are also displayed, **c, d,** Elevated IFNγ levels at the apex of the heart overall (**c,** *, p<0.05 by two-way ANOVA) and in a representative mouse (**d**). **e,** Principal coordinate analysis showing distinct clustering by section, **f, g,** Distinct distribution of

eicosanoid precursors and derivatives at 12 day timepoint in a representative mouse (**f**) and overall (**g**). Non-overlapping notches indicate significantly differing medians (95% confidence), **h, i,** Distribution of purine nucleosides at 12 day timepoint in representative mice (**h**) and overall (**i**). Distinct distribution of purine nucleoside monophosphates and derivatives at 12 day timepoint in a representative mouse (**j**) and overall (**k**).
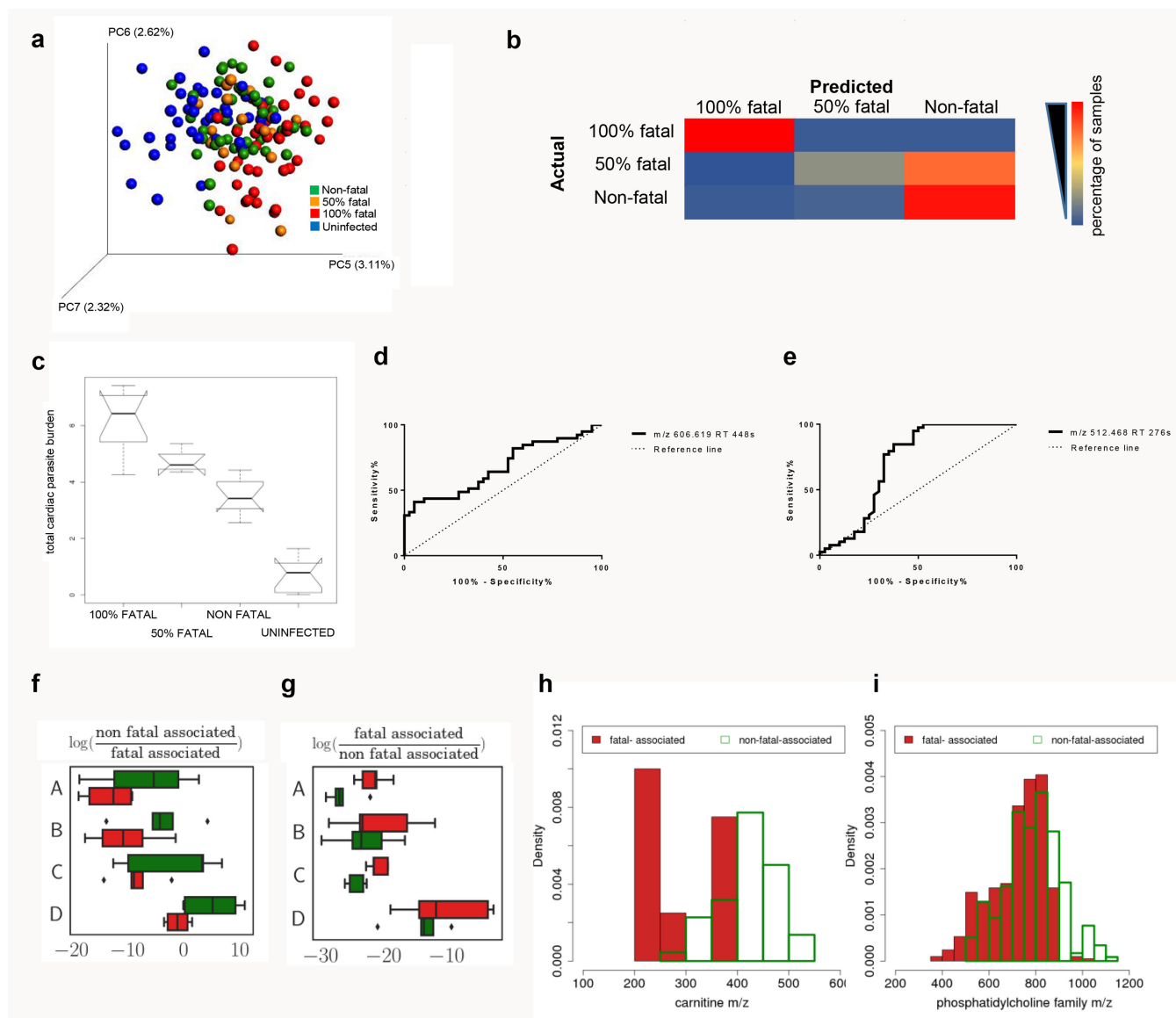
Figure 2. Distinct chemical signatures associated with fatal and non-fatal infection outcomes.
Transverse heart sections were collected 12 days post-infection from two independent
biological experiments. Metabolites were extracted with 3:1 dichloromethane: methanol and
analyzed by LC-MS/MS. Metabolite features for all heart sections from both replicates were
analyzed simultaneously for all statistical testing. **a,** Principal coordinate analysis showing
clustering by infection outcome (p<0.001, Adonis test). Each sphere represents one heart
section from a given mouse. **b,** Random forest classification output showing good
classification accuracy based on metabolite composition for 100% fatal samples and non-
fatal samples. Cells are colored based on the percentage of samples assigned to each
category. Correct classification is along the diagonal. **c,** Total cardiac parasite burden is
associated with infection outcome; non-overlapping notches indicate significantly differing
medians (95% confidence)). **d,** ROC curve for *m/z* 606.619 RT 448s, fatal *vs* non-fatal
groups. p<0.01. **e,** ROC curve for *m/z* 512.468 RT 276s, fatal *vs* non-fatal groups. p<0.005.

f, Boxplots of balance tree analyses (X10 subset) showing a sharp division between fatal and non-fatal samples at the apex of the heart (section position D). **g,** Boxplot of balance tree analyses (differential CL dose subset) showing a sharp division between fatal and non-fatal samples at the base of the heart (section position A). **h,** Differential distribution of detected and identified long and short chain acylcarnitine family members between fatal and non-fatal balances. **i,** Differential mass distribution of detected and identified phosphatidylcholine family members between fatal and non-fatal balances.

Table 1.

Top 30 differential features between heart sections identified via random forest

m/z	RT (s)	Putative annotation
220.119	130	Pantothenic acid
229.155	136	-
240.102	69	-
241.155	26	-
269.212	209	-
276.181	139	Acylcarnitine family member
279.233	223	Linoleic Acid derivative[1]
295.228	173	Linoleic acid derivative[1]
295.228	193	Linoleic acid derivative[1]
301.217	187	-
304.150	80	-
309.207	176	-
317.210	217	11(12)-EpETE derivative[1]
319.227	200	Arachidonic acid methyl ester[1]
319.227	221	Arachidonic acid methyl ester[1]
319.228	175	Arachidonic acid methyl ester[1]
333.205	202	-
335.221	197	O-Arachidonoylglycidol derivative[1]
348.071	29	AMP[2]
362.087	44	Methylated AMP[2]
364.066	31	GMP[2]
371.039	31	-
386.027	26	-
464.082	136	Adenylosuccinic acid[2]
536.185	38	-
557.457	212	-
576.294	163	Phosphatidylcholine family member
611.429	217	-
613.443	224	-
647.450	198	-

[1] Eicosanoid family member or precursor

[2] Nucleosides and derivatives