Plasma mass spectrometry fingerprints induced by bacterial endotoxins in murine models as markers of pathophysiological evolution in sepsis

Author names and affiliations:

Martín Ledesma[1, 7], María Florencia Todero[2], Lautaro Maceira[2], Monica Prieto[3], Carlos Vay[1], Marcelo Galas[4], Beatriz López[5], Noemí Yokobori[6, 7], Bárbara Rearte[2, 7]

1 Laboratorio de Bacteriología, Departamento de Bioquímica Clínica, Hospital de Clínicas José de San Martín; Facultad de Farmacia y Bioquímica, UBA, Buenos Aires, Argentina

2 Instituto de Medicina Experimental (IMEX) – CONICET -Academia Nacional de Medicina, Buenos Aires, Argentina.

3 Servicio Bacteriología Especial. Instituto Nacional de Enfermedades Infecciosas (INEI), ANLIS "Dr. C. G. Malbrán", Buenos Aires, Argentina.

4 Special Program of AMR, Communicable Diseases and Environmental Determinants of Health Department, Panamerican Health Organization, Washington, D.C., USA .

5 Departamento de Bacteriología. Instituto Nacional de Enfermedades Infecciosas (INEI), ANLIS "Dr. C. G. Malbrán", Buenos Aires, Argentina.

6 Instituto Nacional de Enfermedades Infecciosas (INEI), ANLIS "Dr. C. G. Malbrán" Buenos Aires, Argentina.

7. Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

ABSTRACT

Sepsis constitutes a major cause of death in intensive care units. Patients present a dysregulated response to infection, which progresses through different pro/anti-inflammatory phases. The lack of tools to monitor their response constrains the therapeutic approaches. Here, we evaluated a test based on plasma protein fingerprints acquired by MALDI-TOF-mass spectrometry and supervised/unsupervised algorithms to discriminate the different immunological stages of sepsis in lipopolysaccharide-induced murine models with encouraging results. Moreover, our predictive models through machine learning algorithms were able to discriminate the different groups with a sensitivity of up to 95.7% and a specificity of 90.9% depending on the selected peaks number. Potential individual biomarkers associated with each phase were also analysed. Our data reveal the potential of plasma peptidome analysis by MALDI-TOF-mass spectrometry as a highly relevant strategy for sepsis patient stratification that could contribute to therapeutic decisions, depending on the immunological phase that the patient is undergoing.

INTRODUCTION

Sepsis continues to be one the major causes of death in intensive care units (ICU)[1,2]. Estimations indicate a worldwide incidence of 31.5 million cases annually with approximately 7-9 million deaths[3] The precise impact of sepsis in Latin America is elusive due to the limited number of prevalence studies. However, in view of the high incidence rate of infectious diseases along with the inequities in access to health care prevailing there, sepsis can be regarded as an alarming health problem for the region[4-6]. A recent study in Brazil indicates a prevalence of sepsis in ICUs of 30% with a mortality rate of 55%[7], while in Argentina a mortality ranging from of 41.6% to 51%[4] has been ascribed to septic shock.

Considering this critical situation, in its 70th Assembly held in 2017, the World Health Organization reinforced the recommendations to improve the prevention, early diagnosis, follow-up and recording of sepsis cases, along with the promotion of research for the development of new treatments and diagnostic methods[8,9].

Despite the steady advances made in our understanding of sepsis pathophysiology in the last decades, this knowledge has not been fully translated into a substantial improvement in the therapeutic approach. Both diagnosis and treatment of sepsis still present numerous challenges[10]. A comprehensive discussion in 2016 led to the redefinition of sepsis (Sepsis-3) based on consensus recommendations, as a “life-threatening organic dysfunction caused by a dysregulated host response to infection”[11]. This concept highlights the need to integrate the information on the aetiological agent and the host response status, with a strong emphasis in the latter.

In this line, from the immunological point of view, sepsis is characterized by a dysregulated or dysfunctional host immune response to infection that varies over time; clinically, it shows the predominance of an overwhelming pro-inflammatory response during the early stages. As sepsis progresses, an anti-inflammatory and immunosuppressive response becomes predominant, which impacts directly on patient outcome[12] since the most deaths occur during this late stage due to opportunistic reinfections[12-14]. The dynamic nature of sepsis constrains the potential immunotherapy approaches due to the lack of tools to monitor the host immune status promptly.

In spite of the efforts put in biomarker search in the past decades, the application of clinically useful biological molecules or signatures for the diagnosis and the monitoring of sepsis is still finite[15-18]. FDA-approved biomarkers, such as procalcitonin, have a limited application as a guide for antibiotic treatment and as a prognostic marker under certain conditions[16,18-20]. Moreover, none of these approaches focused in the stages of the disease based on its quickly changing immunological status[12-14,16,18-20]. In this light, a fast and readily transferrable test is urgently needed.

Over the past decades, the emergence of "omics" technologies have proven to be a new source of promising results in the study of many diseases including sepsis. Matrix-assisted light desorption ionization-time of flight-mass spectrometry (MALDI-TOF-MS) is a powerful tool for studying protein

molecules present in complex biological samples, with high fidelity and in a few minutes[21]. Notably, this technology proved to be extremely cost-effective in clinical microbiology for the identification of bacteria[22], and it is becoming increasingly available in the clinical diagnosis setting, even in resource-constrained countries. The field of proteome and peptidome analysis of diverse biological fluids has also been dominated by mass spectrometry, and top-down analysis is progressively gaining ground. This approach can provide information about the physiological state of an organism and contribute to disease diagnosis[23], using for this purpose, analysis strategies such as artificial intelligence algorithms. Machine learning (ML) is a concept in which generic algorithms are used to learn about certain sample dataset. The algorithms are trained with these data to look for unique features among groups. This knowledge is then applied to classify or cluster new data and make predictions[24].

Our approach based on the analysis of plasma biomolecular fingerprints has the advantage of allowing the simultaneous evaluation of multiple markers, thus overriding costly and time-consuming purification and immunolabeling steps. In addition, this method would be readily transferrable, considering that MALDI-TOF-MS technology is currently available for microbiological biotyping in several health care facilities and that analysis were performed using the open source software R. Our final goal is to develop a fast, simple and inexpensive tool that can be available at health care centers and can serve as a guide in the management of sepsis.

Around 50% of all sepsis cases are caused by Gram-negative bacteria[25,26] whose membrane lipopolysaccharide (LPS) are one of the main agents associated both inflammatory and anti-inflammatory processes[27,28]. Our group has intensively worked in LPS-based models that emulate these phases[29-35]. Thus, in this study we tested the potential of our approach by developing well defined animal models emulating pro- and anti-inflammatory phases that occur during a septic process in LPS-induced murine models. The final goal of the current work is to develop predictive models based on machine-learning algorithms that would allow the differentiation between the two different phases (pro/anti-inflammatory) of sepsis through the analysis of the mass spectra of plasma specimens generated by MALDI-TOF-MS.

RESULTS

Simulation of pro- and anti-inflammatory sepsis phases

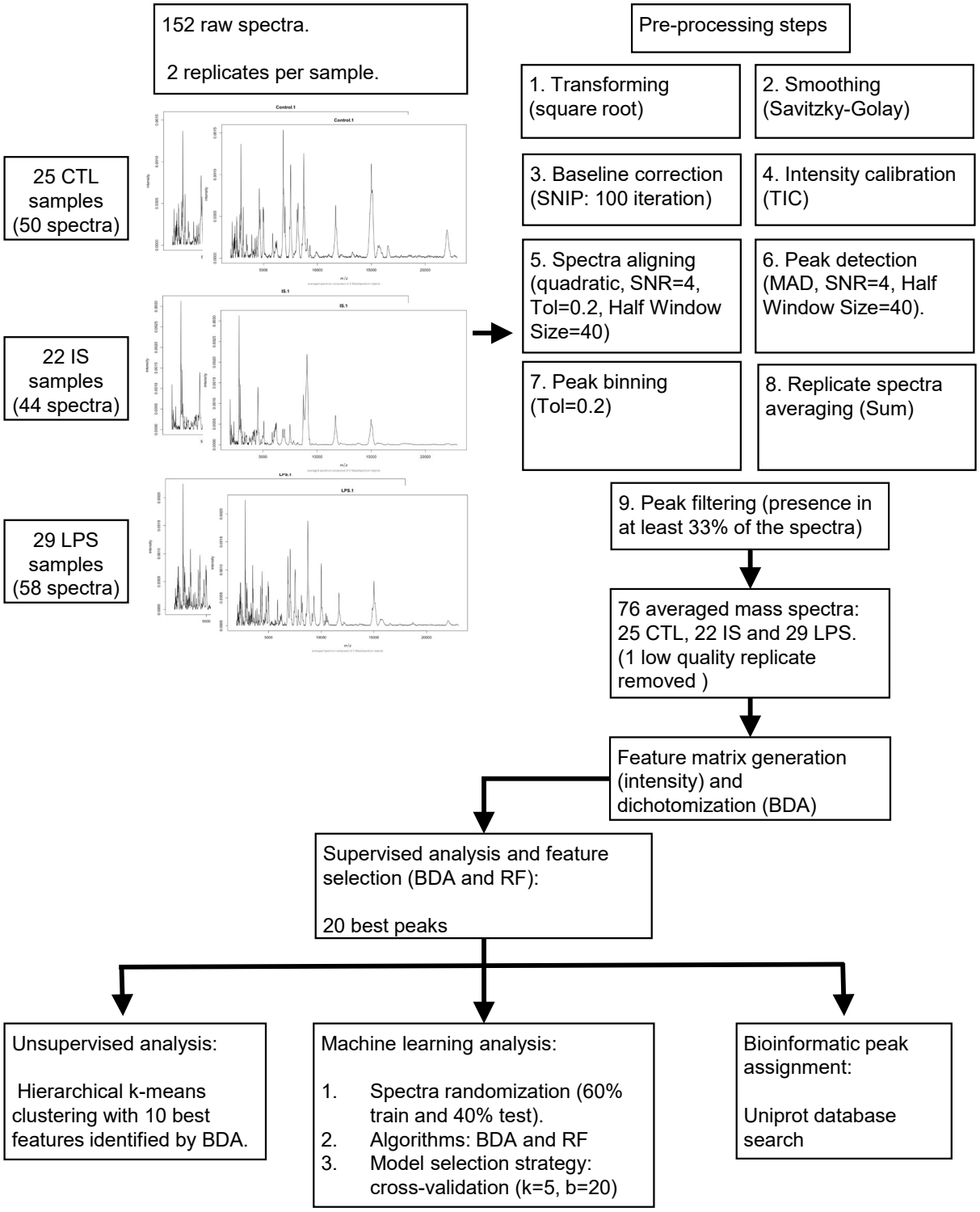
The inflammatory phase was induced with a high dose of LPS (LPS group), whereas the anti inflammatory/immunosuppression phase was reached with an scheme of increasing doses of LPS (IS group) (inoculation and sample collection schemes are described in detail in Suppl. Fig.1). In order to validate the two models, immunological, inflammatory and tissue damage markers were studied in the two experimental groups as well in the control group (CTL group). In accordance with our previous results, the pro-inflammatory phase was characterized by a marked leukopenia (Suppl. Fig 2a), high levels of pro-inflammatory cytokines such as TNF-α in plasma, but also of the anti-inflammatory cytokines IL-10 and TGF-β (Suppl. Fig 3a, b). Tissue enzymes indicative of damage were also elevated (Suppl. Fig 4a-c). On the other hand, in the anti-inflammatory/immunosuppressed (IS) mice,

a leukocytosis due to increased numbers of circulating polymorphonuclear neutrophils (PMN) (Suppl. Fig 2b) along with high plasma levels of TGF- β and low levels of pro-inflammatory cytokines were observed (Suppl. Fig 3c). Tissue damage markers were unaltered in this group (Suppl. Fig 4d-f). As previously described[32,34,35], IS mice had a profound immunological impairment, accompanied by a marked lymphopenia and an increased number of PMN in the spleen, which expressed higher levels of the inhibitory receptor PDL-1 (Suppl. Fig 5a, 5b). Moreover, IS group showed significantly low antibody titres upon immunization with sheep red blood cells (SRBCs) (Suppl. Fig 5c). Collectively, these results indicate that the pro- and anti-inflammatory/immunosuppression phases were emulated in our murine models.

MALDI-TOF MS data analysis

Plasma samples of the different groups (CTL, IS and LPS) were analyzed by MALDI-TOF MS and the spectra were obtained. Representative spectra of each experimental group are shown in Suppl Fig. 6 and the analysis pipeline is detailed in Figure 1. A statistical analysis set was implemented to test if MALDI-TOF MS data were useful to discriminate the three experimental groups. First, a supervised classification algorithm was performed, searching for peaks that best differentiated the CTL, LPS, and IS groups. The 20 best discriminant peaks identified by binary discriminant analysis (BDA) and random forest (RF) algorithms are shown in Suppl. Fig. 7, which partially overlapped. Of the 20 best peaks selected by the BDA algorithm, the 10 most discriminat peaks were chosen for a unsupervised statistical analysis to plot the hierarchical k-means clustering-PCA clusters (Figure 2a). Clusters appeared as three non-superimposed groups, and the first two principal components (PC) explained 76.6% of the variation with good intra cluster homogeneity. Specifically, cluster 2 achieved 100% (16/16) of homogeneity for spectra corresponding to CTL plasma and 64% (16/25) of CTL spectra fell in this cluster (Figure 2b), while for cluster 1 an 85% (29/34) homogeneity was achieved for spectra corresponding to the LPS group and 100% (29/29) of LPS spectra fell in this cluster. Lastly, homogeneity of cluster 3 was 77% (20/26) mostly represented by spectra of the IS group, with 91% of the IS spectra (20/22) included in this cluster. This result show that the immunosuppression/anti-inflammatory and pro-inflammatory phases have distinctive signatures that allow their discrimination.

Figure 1. Analysis Pipeline



Abbreviations. CTL: control group; IS: immunosuppression/anti-inflammatory group; LPS: inflammatory group; BDA: binary discriminant analysis; RF: random forest.

Figure 2

a **b**

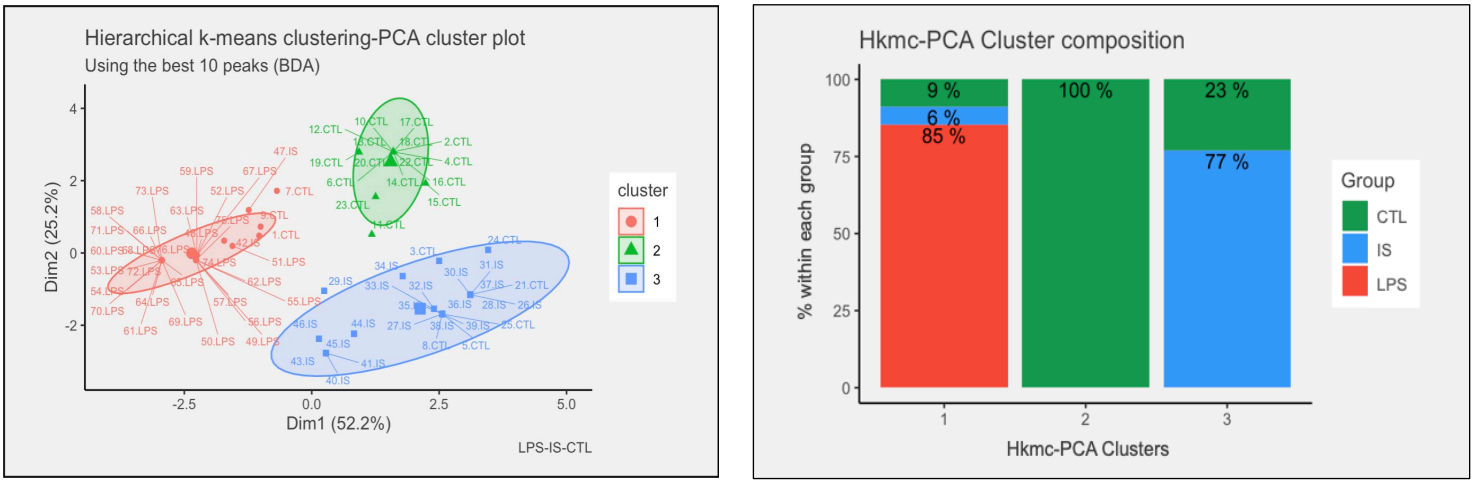


Figure 2. Unsupervised statistical analysis. (a) Hierarchical k-means clustering (Hkmc)-Principal Component Analysis (PCA) cluster plot using the top ten peaks selected by the binary discriminant analysis (BDA) algorithm. Labels contain the mice ID and immunological sepsis status. Principal components 1 (Dim1, x axis) and 2 (Dim2, y axis) are depcted. Spectra were clustered into three groups using the Hierarchical k-means clustering algorithm, which were represented with three different colours. 95% confidence ellipses were added around cluster means, assuming a multivariate normal distribution. (b) Hierarchical k-means clustering-PCA cluster composition. The green color represents the CTL mice group, the blue represents the IS group, and the red color represents the LPS group.

Machine learning analysis

Machine learning (ML) algorithms were used to develop a predictive model based on the plasma fingerprints, using a training data set (14 CTL, 13 IS and 15 LPS samples) challenged with a test data set (11 CTL, 9 IS and 14 LPS samples;. Figure 1). To find the most refined model, a dimensionality reduction was performed through the selection of different numbers of distinctive peaks for each physiological phase to train the algorithms, and the performances on the test set were computed for each condition. Specifically, the top five, ten, fifteen, and twenty discriminant peaks for BDA and RF algorithms were evaluated by cross-validation (CV; 5 folds and 20 repetitions) and percentages of accuracy, sensibility, specificity, negative predictive value (NPV) and positive predictive value (PPV) in the test data set were obtained (Table 1). Among models, RF using the best 20 peaks showed the most consistent values, with sensibility and specificity values above 90%. The BDA model trained with the best five peaks showed the best performance in terms of detection of positive cases, with a sensitivity above 95%; it strategy constitutes a more simple method in terms of both the number of required peaks and from theoretical viewpoint, being an approach more economical in terms of computational demands**.** These results display the potential of plasma fingerprints to find predictive models.

Table 1

Peaks	Algorithm	A (%)	S (%)	E (%)	PPV (%)	NPV (%)
5	BDA	90.6	95.7	80	90.9	89.8
10	BDA	85.6	92	72.3	87.4	81.1
15	BDA	87.6	93.9	74.5	88.5	85.4
20	BDA	90.1	94.3	81.4	91.4	87.3
5	RF	88.2	95.7	72.7	88	88
10	RF	91.2	91.3	90.9	95.5	83.3
15	RF	88.2	91.3	81.8	91.3	81.8
20	RF	94.1	95.7	90.9	95.7	90.9

Table 1. Performance of classification models. Performance values obtained with the top 5, 10, 15, and 20 peaks were presented. A = accuracy, S = sensitivity, E = specificity, PPV = positive predictive value, and NPV = negative predictive value.

263 **Bioinformatic peak assignment**

264 The 20 discriminant peaks (total = 25, 15 accredited by both methods, 5 accredited only by BDA and
265 5 accredited only by RF) were bioinformatically assigned to proteins/peptides matching entries in the
266 Uniprot database according to their m/z listed in Suppl. Table 1. Most hits corresponded to fragments
267 of bigger proteins (Suppl. Table 1). The only peptide identified in its native form was the antimicrobial
268 peptide β-defensin-2, which was selected by both BDA and RF approaches as highly discriminant.

269 Other peptides identified in this analysis were β-enolase (7,643 Da peak), BRCA1 gene-1 protein,
270 and lysosomal α-mannosidase (4,870 Da peak), among others (Suppl. Table 1). These peptides were
271 also found in human plasma by Dufresne et al.[36]. Among fragments of the most abundant plasma
272 proteins decribed by these authors, a fragment that matched the joining region of the immunoglobulin
273 heavy chain was also found in this analysis.
274

275 **DISCUSSION**

276 Sepsis constitutes a highly heterogeneous syndrome of complex pathophysiology. The underlying
277 immunological alterations that include pro-inflammatory, anti-inflammatory and immunosuppressive
278 components[12] are reflected in septic plasma protein composition and varies over time according to
279 disease severity[37]. Thus, plasma represents a favoured biological sample to search for biochemical
280 signals that provide information on the physiological state at a given time. In the last decades, several
281 circulating markers, both soluble as well as cell surface molecules with well recognized functions in
282 sepsis were identified. However, none of these markers, nor a combination of them, showed the
283 specificity and sensitivity required for its clinical use[15,18] .

284 Furthermore, methods that rely on immunolabelling as ELISA or flow cytometry, are highly specific
285 and sensitive for the detection of low abundance markers, but requires costly equipment and
286 consumables, as well as specific technical knowhow. In addition, these techniques are time
287 consuming, limiting their clinical usefulness for early decision making.

288 Regarding biomarker discovery, proteomics proved to be an important source of information.
289 However, in the classical approaches the plasma samples are subjected to depletion of abundant
290 proteins[23] and /or fractionation to identify individual biomarkers[37,38] which requires more analysis
291 time, specific equipment and higher costs.

292 These constraints led us to explore a new top-down strategy that could overcome these difficulties.
293 Thus, having in mind that we aimed to design a readily transferrable method, our approach was meant
294 to be as simple as possible at the beginning, to eventually introduce technical improvements as the
295 analysis demanded. For this reasons pre-processing of the sample was avoided and whole plasma
296 were directly read by MALDI-TOF-MS.

297 Surprisingly, the results described herein showed that the performance of the test with whole plasma
298 and standard experimental settings used in routine microbiological biotyping were excellent in terms

of discriminatory power. First, an ensemble between supervised and unsupervised algorithms allowed the overall discrimination of plasma samples from the three experimental groups, not only the ones representing the immunosuppressive and the inflammatory phases from the control group, but also between them. These results showed that each phase comprise a distinctive MALDI-TOF-MS pattern, which turns this approach into a promising tool to determine different phases in sepsis. A recent study reported a top-down plasma analysis showing that differences between samples from healthy donors and cancer patients could be appreciated without the depletion of large abundant proteins by setting a cutoff at 30KDa[39], supporting our results.

Furthermore, the performance of our machine learning algorithms challenged with the test dataset, was able to identify basal, inflammatory, and immunosuppressive phases in a very satisfactory way, exceeding 90% of accuracy, even supassing 95% in some of the tested coditions. The cross validation strategy used herein allows a fine tuning of the models according to particular needs. Thus depending on the particular clinical and epidemiological settings, algorithms with higher sensitivity, like our 5-peaks BDA model, would be preferred over those with a good overall accuracy. In addition, this model would be more economical in terms of computational demands.

The peaks found in the 2-20KDa spectrum range mostly correspond to small peptides. Plasma peptidome has been proposed as an interesting source of information for diagnostic purposes[36,39]. These peptides are thought to be fragments of larger proteins derived from digestion by proteases or non-enzymatic degradation[40]. Accordingly, recent reports demonstrate the importance of plasma peptidome as a useful tool for clinical monitoring in septic and hemorragic shock[41,42]. Some of the potential peptides identified in our analysis (Suppl. Table 1), were also found by other authors in human plasma[36]. The only peptide assigned in its native form was the β-defensin-2, and was selected by both BDA and RF approaches. This antimicrobial peptide has been proposed as an useful biomarker in sepsis[37]. The identification of our discriminant peaks should be confirmed by more specific methods, and the lack of comprehensive digestome/degradome databases is a well recognized problem in the field[43] that could have conditioned our results.

A major limitation of this study is related to the animal models chosen. Although LPS-induced inflammation/immunosuppression models are widely used, they do not accurately represent the dynamic physiological changes that occur in sepsis[44]. Endotoxin challenge promotes a more quick and transient release of pro/anti-inflammatory mediators compared to the septic processes. However these models allowed us to define two discrete phases that actually occur in a septic process. For this reason, we considered this was the most appropriate approach to test the potential of MALDI-TOF-MS analysis in the proof-of concept phase. We are currently validating our results in a dynamic model of sepsis, the cecal ligation and puncture model, with very encouraging results.

However, our approach has shown numerous advantages. First, MALDI-TOF-MS is a simple and speedy method, which is a critical feature taking into account the dynamism of sepsis. This technology is spreading in the healcare system, particularly in centers were patients with sepsis are usually treated. In addition, we used an accessible sample as plasma, without complex pre-processing steps. In addition, we performed our ML analysis in the open source softwere R. This language, that also

338 supports user friendly web-based interface integration, is another feature that guarantees the
339 transferability even in resource contrained settings.

340 In the light of the results presented herein, we consider that plasma peptidome analysis by MALDI-
341 TOF-MS has the potential to be a highly relevant strategy for sepsis patient stratification. Ideally, it
342 would constitute a powerful tool for therapeutic decision making in sepsis depending on the pro- or
343 anti-inflammatory phase the patient is undergoing.

344

345 MATERIALS AND METHODS

346 Animals

347 Female and male BALB/c mice (8–12 weeks old) were provided by the IMEX-CONICET-Academia
348 Nacional de Medicina, Buenos Aires, Argentina. Animals were maintained under a 12 h light–dark
349 cycle at 22 ± 2 °C and were fed with standard diet and water ad libitum. Animals were bred and
350 housed in accordance with the NIH Guide and Use of Laboratory Animals[45]. Experimental designs
351 were approved by the Committee for the Care and Use of Laboratory Animals (CICUAL) of IMEX-
352 CONICET, Academia Nacional de Medicina de Buenos Aires.

353 Murine Models

354 Inoculation schemes are depicted in Supplementary Figure 1.

355 *Inflammatory phase model (LPS group):* BALB/c mice were inoculated with an LD_{50} of
356 lipopolysaccharide of *Escherichia coli* O111: B4 (LPS) (100 ug/ mouse) (Sigma-Aldrich, St Louis, MO,
357 USA) intraperitoneally (i. p.)[29]. Plasma samples were obtained 1.5 h and 6 h after the injection.

358 *Anti-inflammatory phase model (IS group):* BALB/c mice were daily inoculated with LPS for 10
359 consecutive days. The inoculation scheme consisted of 5 µg/mouse i.p. for the first 4 days, followed
360 by 50 µg/mouse i.p. for 3 days, and then 100 µg/mouse i.p. for the last 3 days[34]. 24 h after the last
361 LPS dose, plasma samples were obtained.

362 A third group of mice inoculated with vehicle (saline solution) served as control *(CTL group)*. The
363 plasma in this group was collected at the same time as indicated in the experimental groups .

364 The submandibular bleeding method was used to obtain plasma in order to maximize the quality of
365 the sample. The heparinized samples were centrifuged twice (400xg, 10 min at 4ºC) and plasma were
366 stored at -20ºC until analysis.

367 Immunological, hematological and biochemical parameters

368 *Peripheral blood leukocytes count*: blood cells obtained by a submandibular puncture in heparinized
369 tubes were analyzed by Coulter hematology analyzer (Diatron Abacus Junior Vet, Budapest,
370 Hungary) at 1.5 h and 24 h post LPS for the LPS and the IS groups respectively. *Cytokine*
371 *determination:* pro-inflammatory (TNF-α; IL-12p70; IFN-γ; IL-6) and anti-inflammatory cytokines (IL-
372 10, TGF-β) were determined in plasma by ELISA at the indicated time points (OptEIA set; BD
373 Biosciences, San Diego, CA, USA) according to the manufacturer's instructions. *Tissue damage*
374 *markers:* creatinine (Cre), alanine aminotransferase (ALT) and aspartate transaminase (AST) levels
375 were determined in plasma, at 6 and 24 h in the LPS and IS groups respectively, using a kit from
376 BioTecnica (Varginha, Minas Gerais, Brazil) and the MINDRAY BS-200E autoanalyzer according to
377 the manufacturer's instructions. *Flow cytometry*: twenty four hours after the last LPS dose splenocytes
378 suspension was obtained and the cells were immunolabeled as described previously[34,35] for the
379 following cell types: T lymphocytes (FITC)-CD4 (clone RM4-5) and (PECy5)-CD8 (clone 53-6.7), B
380 lymphocytes (FITC)-CD45R-B220 (clone RA3-6B2), neutrophils (FITC)-CD11b (clone M1/70)/ (PE)-
381 Ly6G clone 1A8), and the expression levels of (PE)-PDL-1 (clone MIH5) on CD11b gate of myeloid
382 lineage cells. Antibodies were obtained from Invitrogen and BD Pharmingen™ . The acquisition was
383 carried out by a Becton Dickinson FACScan using Cell Quest software (Becton Dickinson, San Jose,
384 CA,USA). *Antibodies titre:* primary antibody response to sheep red blood cells (SRBC) by immunizing
385 animals 24 h after the last dose of LPS (SRBC: $5x10^8$/mouse, 0.1ml i.p.). The anti-SRBC antibody
386 titre was evaluated in serum through a hemagglutination assay seven days after immunization as
387 described previously[34,35].

388 **Acquisition of MALI-TOF-MS spectra**

389 Plasma samples were analyzed with the MALDI Biotyper System (Bruker Daltonik GmbH, Bremen,
390 Germany). The total number of plasma samples analyzed per group were the following: LPS
391 group=n:29; IS group=n: 22 ; CTL group=n: 25. The standard method used for routine microbial
392 biotyping was used to analyse the plasma samples. Briefly, one µl of plasma was loaded onto each
393 spot of the metal plate in duplicate. After the sample was allowed to dry, one µl of the matrix (alpha-
394 cyano-4-hydroxycinnamic acid matrix in 50% acetonitrile and 2.5% trifluoroacetic acid) was added to
395 each spot and was allowed to dry again. Duplicates were acquired seeding two spots per sample.

396 Continuous mass spectra were obtained with a Microflex LT/SH MALDI-TOF mass spectrometer (with
397 flexControl software version 3.4.135.0; ionization mode: LD+; acquisition method: MBT_FC.par;
398 acquisition mode: qsim; tofMode: linear; acquisition Operator Mode: linear; digitizerType: Bruker
399 BD0G5; instrument ID: 269944.00510; instrument type: Microflex) within a mass range of 2,000-
400 20,000 Da. Spectra were obtained in the manual mode, using 60% of laser intensity with 40 laser
401 repetition in each shot and reaching between 400-500 spectra by acquisition. The rationale of
402 choosing these acquisition conditions were 1) to use the acquisition range set for bacterial biotyping
403 to avoid a specific calibration step and 2) to avoid the >20,000 Da range where highly abundant
404 proteins such as albumin mask other proteins with lower abundance. Spectra were internally

405 calibrated and controlled every day using *Escherichia coli* ribosomal proteins (bacterial test standard;
406 Bruker Daltonik GmbH).

407 **MALDI-TOF-MS data pre-processing**

408 We generated a dataset containing 152 mass spectra from mice plasma samples (50 belonged to the
409 CTL group, 44 to the IS group and 58 to the LPS group), including duplicates. Mass spectra were
410 read as fid/aqus files with MALDIquantForeign (v0.10)[46] and they were processed using MALDIquant
411 (v1.16.2)[47] R package. Briefly, spectra were square root-transformed and smoothed using the
412 Savitzky-Golay algorithm. Spectra were baseline-corrected, applying the SNIP process across 100
413 iterations. The peak detection was carried out by a function that estimates the noise of mass
414 spectrometry data by calculating the median absolute deviation (MAD). The *s*ignal-to-*n*oise-*r*atio
415 (SNR) was set up in 4, with a halfWindowSize of 40 and a tolerance of 0.2 for peak binning. One
416 replicate corresponding to the CTL group was removed due to low quality. After averaging the
417 replicates, 76 averaged mass spectra were obtained, of which 25 belonged to the Control group, 22
418 to the IS group and 29 to LPS group. Peaks that occured in less than 33% of spectra were removed.
419 MALDI-TOF-MS data were transformed by a categorization of the peak intensity, performed with the
420 Binda R package[48]. The algorithm compares the intensity value with the peak group average,
421 returning 1 when it was equal or higher and 0 when it was lower. The analysis workflow is depicted in
422 the Figure 1.

423 **Statistical analysis**

424 *Immunological, hematological and biochemical parameters.* Statistical analysis and plotting were
425 achieved using Graph Pad Prism 6 software (GraphPad Software, San Diego, CA). Values are
426 expressed as the mean ± standard error of the mean (SEM) of n samples. The number of samples
427 (n) is refers to number of mice or biological repeats analyzed each time. These biological repeats
428 were used for the statistical analysis. No outliers were removed. The assumption test to determine
429 the Gaussian distribution was performed by the Kolmogorov and Smirnov method. For parameters
430 with a Gaussian distribution, the differences between 2 experimental groups were assessed by
431 unpaired Student's t test. For multiple group comparisons, as the peripheral blood leukocytes
432 differences, were analyzed using one-way ANOVA followed by a Tukey's multiple comparison test.
433 For parameters with a non-Gaussian distribution, comparisons between 2 experimental groups were
434 performed using Mann–Whitney *U* test. All statistical tests were interpreted in a two-tailed fashion and
435 a $P<0.05$ are considered statistically significant; the level of significance is indicated by asterisk
436 numbers as follows: *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$, ****, $P<0.0001$.The statistical analysis used
437 as also the number of samples (n) and the repetitions of each experimental design are reported in
438 each figure legend.

MALDI-TOF-MS data. Statistical analysis and plotting were performed using Rstudio Software. Programmed peaks selection was performed in the entire dataset seeking for biomarkers of each class representing the different sepsis phases (LPS and IS groups) by the Binary Discriminant Analysis (BDA) algorithm. The algorithm outputs the t.score (Class means vs. Pooled mean) of each peak. The sign of the t.score indicates the presence (positive t.score) or absence (negative t.score) of that peak in each group. A significance level of 95% was achieved if the t.score was equal or higher than 2.5 and equal or less than -2.5. The best-extracted features were then used to perform a hierarchical k-means clustering-principal component analysis (PCA) with the factoextra R package[49]. The binary distance was used to measure dissimilarity between observations, ward.D2 was the agglomeration method used, and k was set to 4. In addition, spectra were analyzed with the random forest (RF) classifier. Afterwards, two machine learning (ML) methods were applied in the dataset to train two classifiers: BDA and random forest RF[50] also with RStudio software.

These classification models, used for the feature extraction and ranking, performe their selection based on distinctive criteria. For instance, BDA seeks for differences by measuring the t-scores between the class mean and pooled mean, while RF computes the total decrease in node impurities from splitting on the variable, averaged over all decisition trees, and the Gini index measures the node impurity. Initially, the dataset was randomly partitioned into a training set (60% of plasma samples) and a test set (40% of plasma samples). Programmed feature (peaks) selection was performed with the respective algorithms in the training set seeking for discriminant peaks corresponding to each sepsis phase. The extracted features were then used to train several ML models. Metrics such as accuracy, sensitivity, and specificity, were used to evaluate the performances of all resulting models in the test set by a cross-validation strategy. The entire dataset analyzed in this work together with the R sample code used for the analysis is available at: https://github.com/MarManLed/SepsisData.

Bioinformatic peak assignment

Uniprot database was used to assign the best 20 discriminant peaks. The search was performed on the *Mus musculus* data base (ID) and within a 15 Da interval around each peak.

Acknowledgements

We thank Dr. Martín Isturiz for his eternal accompaniment and support for the development of this study. His differential contribution stems from his singular thought. Forever in our hearts, Dr. Isturiz. We thank Dr. Viviana Ritacco for her thoughtful comments on the manuscript.

Funding

This study was supported by the Agencia Nacional de Promoción Científica y Tecnológica [grant number PICT 2015-0412] Buenos Aires, Argentina.

Author contribution

ML designed and performed the bioinformatic analysis. BR conceived and designed the experiments. MFT, LM and BR performed the experiments with mice and analyzed data. MP and CV provided materials and technical support. MG and BL provided materials and collaborated in study design. ML and NY contributed in data curation. NY and BR analyzed, discussed and realized the study. BR was responsible for funding acquisition and study supervision. Manuscript was written by NY and BR, and was corrected by ML and BL. All the authors revised and approved the final version of the manuscript.

Competing interests

The authors declare that no competing interests exist.

References

1 Novosad, S. A., Sapiano, M. R., Grigg, C., Lake, J., Robyn, M., Dumyati, G., Felsen, C., Blog, D., Dufort, E., Zansky, S., Wiedeman, K., Avery, L., Dantes, R. B., Jernigan, J. A., Magill, S. S., Fiore, A. & Epstein, L. Vital Signs: Epidemiology of Sepsis: Prevalence of Health Care Factors and Opportunities for Prevention. *MMWR Morb Mortal Wkly Rep* **65**, 864-869, doi:10.15585/mmwr.mm6533e1 (2016).

2 Vincent, J. L., Rello, J., Marshall, J., Silva, E., Anzueto, A., Martin, C. D., Moreno, R., Lipman, J., Gomersall, C., Sakr, Y. & Reinhart, K. International study of the prevalence and outcomes of infection in intensive care units. *JAMA* **302**, 2323-2329, doi:10.1001/jama.2009.1754 (2009).

3 Fleischmann, C., Scherag, A., Adhikari, N. K., Hartog, C. S., Tsaganos, T., Schlattmann, P., Angus, D. C. & Reinhart, K. Assessment of Global Incidence and Mortality of Hospital-treated Sepsis. Current Estimates and Limitations. *Am J Respir Crit Care Med* **193**, 259-272, doi:10.1164/rccm.201504-0781OC (2016).

4 Estenssoro, E., Kanoore Edul, V. S., Loudet, C. I., Osatnik, J., Rios, F. G., Vazquez, D. N., Pozo, M. O., Lattanzio, B., Palizas, F., Klein, F., Piezny, D., Rubatto Birri, P. N., Tuhay, G., Diaz, A., Santamaria, A., Zakalik, G. & Dubin, A. Predictive Validity of Sepsis-3 Definitions and Sepsis Outcomes in Critically Ill Patients: A Cohort Study in 49 ICUs in Argentina. *Crit Care Med* **46**, 1276-1283, doi:10.1097/CCM.0000000000003208 (2018).

5 Estenssoro, E., Loudet, C. I., Edul, V. S. K., Osatnik, J., Rios, F. G., Vasquez, D. N., Pozo, M. O., Lattanzio, B., Palizas, F., Klein, F., Piezny, D., Rubatto Birri, P. N., Tuhay, G., Diaz, A., Santamaria, A., Zakalik, G. & Dubin, A. Health inequities in the diagnosis and outcome of sepsis in Argentina: a prospective cohort study. *Crit Care* **23**, 250, doi:10.1186/s13054-019-2522-6 (2019).

6 Azevedo, L. C. P., Cavalcanti, A. B., Lisboa, T., Pizzol, F. D. & Machado, F. R. Sepsis is an important healthcare burden in Latin America: a call to action! *Rev Bras Ter Intensiva* **30**, 402-404, doi:10.5935/0103-507X.20180061 (2018).

7 Machado, F. R., Cavalcanti, A. B., Bozza, F. A., Ferreira, E. M., Angotti Carrara, F. S., Sousa, J. L., Caixeta, N., Salomao, R., Angus, D. C. & Pontes Azevedo, L. C. The epidemiology of sepsis in Brazilian intensive care units (the Sepsis PREvalence Assessment Database, SPREAD): an observational study. *Lancet Infect Dis* **17**, 1180-1189, doi: 10.1016/S1473-3099(17)30322-5 (2017).

8 Organization, W. H. WHO: WHA70.7 Improving the Prevention, Diagnosis and Clinical Management of Sepsis. doi:http://apps.who.int/gb/e/e_wha70.html. Accessed October 2, 2017 (2017).

9 Reinhart, K., Daniels, R., Kissoon, N., Machado, F. R., Schachter, R. D. & Finfer, S. Recognizing Sepsis as a Global Health Priority - A WHO Resolution. *N Engl J Med* **377**, 414-417, doi:10.1056/NEJMp1707170 (2017).

10 Rhodes, A., Evans, L. E., Alhazzani, W., Levy, M. M., Antonelli, M., Ferrer, R., Kumar, A., Sevransky, J. E., Sprung, C. L., Nunnally, M. E., Rochwerg, B., Rubenfeld, G. D., Angus, D. C., Annane, D., Beale, R. J., Bellinghan, G. J., Bernard, G. R., Chiche, J. D., Coopersmith, C., De Backer, D. P., French, C. J., Fujishima, S., Gerlach, H., Hidalgo, J. L., Hollenberg, S. M., Jones, A. E., Karnad, D. R., Kleinpell, R. M., Koh, Y., Lisboa, T. C., Machado, F. R., Marini, J. J., Marshall, J. C., Mazuski, J. E., McIntyre, L. A., McLean, A. S., Mehta, S., Moreno, R. P., Myburgh, J., Navalesi, P., Nishida, O., Osborn, T. M., Perner, A., Plunkett, C. M., Ranieri, M., Schorr, C. A., Seckel, M. A., Seymour, C. W., Shieh, L., Shukri, K. A., Simpson, S. Q., Singer, M., Thompson, B. T., Townsend, S. R., Van der Poll, T., Vincent, J. L., Wiersinga, W. J., Zimmerman, J. L. & Dellinger, R. P. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive Care Med* **43**, 304-377, doi:10.1007/s00134-017-4683-6 (2017).

11 Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J. D., Coopersmith, C. M., Hotchkiss, R. S., Levy, M. M., Marshall, J. C., Martin, G. S., Opal, S. M., Rubenfeld, G. D., van der Poll, T., Vincent, J. L. & Angus, D. C. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* **315**, 801-810, doi: 10.1001/jama.2016.0287 (2016).

12 Rubio, I., Osuchowski, M. F., Shankar-Hari, M., Skirecki, T., Winkler, M. S., Lachmann, G., La Rosee, P., Monneret, G., Venet, F., Bauer, M., Brunkhorst, F. M., Kox, M., Cavaillon, J. M., Uhle, F., Weigand, M. A., Flohe, S. B., Wiersinga, W. J., Martin-Fernandez, M., Almansa, R., Martin-Loeches, I., Torres, A., Giamarellos-Bourboulis, E. J., Girardis, M., Cossarizza, A., Netea, M. G., van der Poll, T., Scherag, A., Meisel, C., Schefold, J. C. & Bermejo-Martin, J. F. Current gaps in sepsis immunology: new opportunities for translational research. *Lancet Infect Dis*, doi: 10.1016/S1473-3099(19)30567-5 (2019).

13 van der Poll, T., van de Veerdonk, F. L., Scicluna, B. P. & Netea, M. G. The immunopathology of sepsis and potential therapeutic targets. *Nat Rev Immunol* **17**, 407-420, doi: 10.1038/nri.2017.36 (2017).

14 Venet, F., Rimmele, T. & Monneret, G. Management of Sepsis-Induced Immunosuppression. *Crit Care Clin* **34**, 97-106, doi: 10.1016/j.ccc.2017.08.007 (2018).

15 Parlato, M., Philippart, F., Rouquette, A., Moucadel, V., Puchois, V., Blein, S., Bedos, J. P., Diehl, J. L., Hamzaoui, O., Annane, D., Journois, D., Ben Boutieb, M., Esteve, L., Fitting, C., Treluyer, J. M., Pachot, A., Adib-Conquy, M., Cavaillon, J. M., Misset, B. & Captain Study, G. Circulating biomarkers may be unable to detect infection at the early phase of sepsis in ICU patients: the CAPTAIN prospective multicenter cohort study. *Intensive Care Med* **44**, 1061-1070, doi:10.1007/s00134-018-5228-3 (2018).

16 Gunsolus, I. L., Sweeney, T. E., Liesenfeld, O. & Ledeboer, N. A. Diagnosing and Managing Sepsis by Probing the Host Response to Infection: Advances, Opportunities, and Challenges. *J Clin Microbiol* **57**, doi: 10.1128/JCM.00425-19 (2019).

17 Pierrakos, C. & Vincent, J. L. Sepsis biomarkers: a review. *Crit Care* **14**, R15, doi: 10.1186/cc8872 (2010).

18 van Engelen, T. S. R., Wiersinga, W. J., Scicluna, B. P. & van der Poll, T. Biomarkers in Sepsis. *Crit Care Clin* **34**, 139-152, doi: 10.1016/j.ccc.2017.08.010 (2018).

554 19 Al Jalbout, N., Troncoso, R., Jr., Evans, J. D., Rothman, R. E. & Hinson, J. S. Biomarkers and Molecular Diagnostics
555 for Early Detection and Targeted Management of Sepsis and Septic Shock in the Emergency Department. *J Appl*
556 *Lab Med* **3**, 724-729, doi: 10.1373/jalm.2018.027425 (2019).
557
558 20 Schenz, J., Weigand, M. A. & Uhle, F. Molecular and biomarker-based diagnostics in early sepsis: current
559 challenges and future perspectives. *Expert Rev Mol Diagn*, 1-10, doi:10.1080/14737159.2020.1680285 (2019).
560
561 21 Angeletti, S. Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical
562 microbiology. *J Microbiol Methods* **138**, 20-29, doi: 10.1016/j.mimet.2016.09.003 (2017).

563 22 Hou, T. Y., Chiang-Ni, C. & Teng, S. H. Current status of MALDI-TOF mass spectrometry in clinical microbiology. *J*
564 *Food Drug Anal* **27**, 404-414, doi: 10.1016/j.jfda.2019.01.001 (2019).

565 23 Ludwig, K. R. & Hummon, A. B. Mass spectrometry for the discovery of biomarkers of sepsis. *Mol Biosyst* **13**,
566 648-664, doi:10.1039/c6mb00656f (2017).
567
568 24 López Fernández, H., Reboiro-Jato, M., Pérez Rodríguez, J. A., Fdez-Riverola, F. & Glez-Peña, D. Implementing
569 effective machine learning-based workflows for the analysis of mass spectrometry data. *Journal of Integrated*
570 *OMICS* **6**, doi:10.5584/jiomics.v6i1.196 (2016).
571
572 25 Dolin, H. H., Papadimos, T. J., Chen, X. & Pan, Z. K. Characterization of Pathogenic Sepsis Etiologies and Patient
573 Profiles: A Novel Approach to Triage and Treatment. *Microbiol Insights* **12**, doi:10.1177/1178636118825081
574 (2019).
575
576 26 van Engelen, T. S. R., Joost Wiersinga, W. & van der Poll, T. Phatogenesis of sepsis. *Handbook of Sepsis* Ch.
577 Chapter 3, 31-43 doi: 10.1007/978-3-319-73506-1_3 (2018).
578
579 27 Genga, K. R., Shimada, T., Boyd, J. H., Walley, K. R. & Russell, J. A. The Understanding and Management of
580 Organism Toxicity in Septic Shock. *J Innate Immun* **10**, 502-514, doi:10.1159/000487818 (2018).
581
582 28 Opal, S. M., Scannon, P. J., Vincent, J. L., White, M., Carroll, S. F., Palardy, J. E., Parejo, N. A., Pribble, J. P. &
583 Lemke, J. H. Relationship between plasma levels of lipopolysaccharide (LPS) and LPS-binding protein in patients
584 with severe sepsis and septic shock. *J Infect Dis* **180**, 1584-1589, doi:10.1086/315093 (1999).
585
586 29 Cordoba-Moreno, M. O., Todero, M. F., Fontanals, A., Pineda, G., Daniela, M., Yokobori, N., Ramos, M. V.,
587 Barrientos, G., Toblli, J. E., Isturiz, M. A. & Rearte, B. Consequences of the Lack of IL-10 in Different Endotoxin
588 Effects and its Relationship With Glucocorticoids. *Shock* **52**, 264-273, doi:10.1097/SHK.0000000000001233
589 (2019).
590
591 30 Landoni, V. I., Chiarella, P., Martire-Greco, D., Schierloh, P., van-Rooijen, N., Rearte, B., Palermo, M. S., Isturiz,
592 M. A. & Fernandez, G. C. Tolerance to lipopolysaccharide promotes an enhanced neutrophil extracellular traps
593 formation leading to a more efficient bacterial clearance in mice. *Clin Exp Immunol* **168**, 153-163,
594 doi:10.1111/j.1365-2249.2012.04560.x (2012).
595
596 31 Martire-Greco, D., Landoni, V. I., Chiarella, P., Rodriguez-Rodrigues, N., Schierloh, P., Rearte, B., Isturiz, M. A. &
597 Fernandez, G. C. all-trans-retinoic acid improves immunocompetence in a murine model of lipopolysaccharide-
598 induced immunosuppression. *Clin Sci (Lond)* **126**, 355-365, doi:10.1042/CS20130236 (2014).
599

600 32 Montagna, D. R., Duarte, A., Todero, M. F., Ruggiero, R. A., Isturiz, M. & Rearte, B. Meta-tyrosine modulates the
601 immune response induced by bacterial endotoxins. *Immunobiology* **225**, 151856,
602 doi:10.1016/j.imbio.2019.10.005 (2020).
603
604 33 Rearte, B., Landoni, V., Laborde, E., Fernandez, G. & Isturiz, M. Differential effects of glucocorticoids in the
605 establishment and maintenance of endotoxin tolerance. *Clin Exp Immunol* **159**, 208-216, doi:10.1111/j.1365-
606 2249.2009.04052.x (2010).
607
608 34 Rearte, B., Maglioco, A., Balboa, L., Bruzzo, J., Landoni, V. I., Laborde, E. A., Chiarella, P., Ruggiero, R. A.,
609 Fernandez, G. C. & Isturiz, M. A. Mifepristone (RU486) restores humoral and T cell-mediated immune response
610 in endotoxin immunosuppressed mice. *Clin Exp Immunol* **162**, 568-577, doi:10.1111/j.1365-2249.2010.04262.x
611 (2010).
612
613 35 Rearte, B., Maglioco, A., Machuca, D., Greco, D. M., Landoni, V. I., Rodriguez-Rodrigues, N., Meiss, R., Fernandez,
614 G. C. & Isturiz, M. A. Dehydroepiandrosterone and metyrapone partially restore the adaptive humoral and
615 cellular immune response in endotoxin immunosuppressed mice. *Innate Immun* **20**, 585-597,
616 doi:10.1177/1753425913502243 (2014).
617
618 36 Dufresne, J., Bowden, P., Thavarajah, T., Florentinus-Mefailoski, A., Chen, Z. Z., Tucholska, M., Norzin, T., Ho, M.
619 T., Phan, M., Mohamed, N., Ravandi, A., Stanton, E., Slutsky, A. S., Dos Santos, C. C., Romaschin, A., Marshall, J.
620 C., Addison, C., Malone, S., Heyland, D., Scheltens, P., Killestein, J., Teunissen, C., Diamandis, E. P., Siu, K. W. M.
621 & Marshall, J. G. The plasma peptidome. *Clin Proteomics* **15**, 39, doi:10.1186/s12014-018-9211-3 (2018).
622
623 37 Hayashi, N., Yamaguchi, S., Rodenburg, F., Ying Wong, S., Ujimoto, K., Miki, T. & Iba, T. Multiple biomarkers of
624 sepsis identified by novel time-lapse proteomics of patient serum. *PLoS One* **14**, e0222403,
625 doi:10.1371/journal.pone.0222403 (2019).
626
627 38 Harberts, E., Liang, T., Yoon, S. H., Opene, B. N., McFarland, M. A., Goodlett, D. R. & Ernst, R. K. Toll-like Receptor
628 4-Independent Effects of Lipopolysaccharide Identified Using Longitudinal Serum Proteomics. *J Proteome Res*
629 **19**, 1258-1266, doi:10.1021/acs.jproteome.9b00765 (2020).
630
631 39 Cheon, D. H., Nam, E. J., Park, K. H., Woo, S. J., Lee, H. J., Kim, H. C., Yang, E. G., Lee, C. & Lee, J. E. Comprehensive
632 Analysis of Low-Molecular-Weight Human Plasma Proteome Using Top-Down Mass Spectrometry. *J Proteome*
633 *Res* **15**, 229-244, doi:10.1021/acs.jproteome.5b00773 (2016).
634
635 40 Greening, D. W. & Simpson, R. J. Characterization of the Low-Molecular-Weight Human Plasma Peptidome.
636 *Methods Mol Biol* **1619**, 63-79, doi:10.1007/978-1-4939-7057-5_6 (2017).
637
638 41 Bauza-Martinez, J., Aletti, F., Pinto, B. B., Ribas, V., Odena, M. A., Diaz, R., Romay, E., Ferrer, R., Kistler, E. B.,
639 Tedeschi, G., Schmid-Schonbein, G. W., Herpain, A., Bendjelid, K. & de Oliveira, E. Proteolysis in septic shock
640 patients: plasma peptidomic patterns are associated with mortality. *Br J Anaesth* **121**, 1065-1074,
641 doi:10.1016/j.bja.2018.05.072 (2018).
642
643 42 Aletti, F., Maffioli, E., Negri, A., Santamaria, M. H., DeLano, F. A., Kistler, E. B., Schmid-Schonbein, G. W. &
644 Tedeschi, G. Peptidomic Analysis of Rat Plasma: Proteolysis in Hemorrhagic Shock. *Shock* **45**, 540-554,
645 doi:10.1097/SHK.0000000000000532 (2016).
646

43 Shen, Y., Liu, T., Tolić, N., Petritis, B. O., Zhao, R., Moore, R. J., Purvine, S. O., Camp, D. G. & Smith, R. D. Strategy for Degradomic-Peptidomic Analysis of Human Blood Plasma. *Journal of Proteome Research* **9**, 2339-2346, doi:10.1021/pr901083m (2010).

44 Lewis, A. J., Seymour, C. W. & Rosengart, M. R. Current Murine Models of Sepsis. *Surg Infect (Larchmt)* **17**, 385-393, doi:10.1089/sur.2016.021 (2016).

45 Institute of Laboratory Animal Resources (U.S.). Committee on Care and Use of Laboratory Animals and National Institutes of Health (U.S.). Division of Research Resources. Guide for the care and use of laboratory animals. Bethesda, MD: US Department Health and Human Services, Public Health Service, National Institutes of Health.. (1985).

46 MALDIquantForeign: Import/Export Routines for MALDIquant. v. 0.12 (2019).

47 Gibb, S. & Strimmer, K. MALDIquant: a versatile R package for the analysis of mass spectrometry data. *Bioinformatics* **28**, 2270-2271, doi:10.1093/bioinformatics/bts447 (2012).

48 Gibb, S., Strimmer, K. binda: Multi-Class Discriminant Analysis using Binary Predictors. R package version 1.0.3. (2015).

49 factoextra: Extract and Visualize the Results of Multivariate Data Analyses v. R package version 1.0.6 (2019).

50 Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2**, 18-22, doi: 10.4236/ijcns.2016.95010 (2002).