Yellow fever disease severity is driven by an acute cytokine storm modulated by an interplay between the human gut microbiome and the metabolome.

Adam-Nicolas Pelletier, Mateus Tomazella, Karina de Carvalho, Andre Nicolau, Marianna Marmoratto, Cassia Silveira, Jorge Khalil, Helder Nakaya, Esper Kallas, Michael Diamond, Rafick-Pierre Sekaly

Abstract

Infection with Yellow Fever (YF) can lead to multiple outcomes ranging from death from total organ failure to clearance of viremia and survival. The mechanisms underlying these differences in clinical outcomes have yet to be defined. We had access to a cohort of YF infected subjects that showed these range of outcomes. An unbiased integrated OMICs approach was used to identify pathways and effector molecules that drive the severe disease and death as compared to resolution of infection. We used the MELD and SIC score as objective markers of disease severity. We show that a specific signature of upregulated innate pro-inflammatory cytokines significantly demarcates subjects with severe disease leading to death from subjects who clear virus. Pathogen sequencing showed heightened levels of proteolytic bacteria at the i.e Actinobacteria and these were correlated to lower levels of tryptophan and tyrosine amino acids measured by untargeted metabolomics. These two features were significantly associated to MELD scores synonymous of milder disease. Propionate a bacterial metabolite that triggers Treg differentiation that can as well limit the hyperimmune activation associated to severe outcome was also associated to improved outcome. Our results suggest a model whereby
proteolytic bacteria limit the availability of the aromatic amino acid pool available for cytokine production thereby preventing the induction of the cytokine storm that is associated to severe disease and death.

INTRODUCTION

Yellow fever virus (YFV) infection is an acute hemorrhagic disease transmitted by Aedes aegypti. Outbreaks of YFV constitute a growing concern, aggravated by the adaptation of Aedes aegypti to urban environments and dwindling immunization coverage. In its severe clinical manifestation, yellow fever (YF) leads to sepsis, fever, hepatitis, renal failure, coagulopathy, hemorrhage, shock and 20-50% lethality. There are no specific antiviral drugs available to treat YFV infection, and multiorgan failure is managed with supportive care.

The mechanisms leading up to the hemorrhagic fever have yet to be elucidated. In vitro studies with hepatocyte infection have shown a substantial induction of proinflammatory responses, including IL-1β, IL-4, IL-6, IL-8, CCL5, IL-10 and TNFa, consistent with a cytokine storm syndrome (CSS). This reflects with observations from infection with another member of the Flavivirus family, Dengue virus, which is known to trigger a CSS in the 10-25% of infected patients who progress to Dengue Hemorrhagic Fever. More recently, SARS-CoV2 infection has also been known to induce a CSS culminating in multi-organ failure.
What remains unclear is why only some individuals mount a CSS response. Prior observations in the context of YFV infection has associated a higher frequency of band neutrophils (immature) in patients with a lethal outcome, which is also reminiscent of SARS-CoV2 infection. One suggested mechanism has been an increase in the release of Neutrophil Extracellular Traps (NETs) through NETosis, which then act as Damage-Associated Molecular Patterns (DAMP) to further cause inflammation when produced in excess.

Growing evidence suggests that environmental factors at the time of infection can influence immune response to antigens (from pathogens or vaccination). We have previously shown that higher levels of activated and proinflammatory immune cell subsets as well as lymph node inflammation and fibrosis in Ugandan cohorts correlated with lower neutralizing Ab titers after YFV-17D vaccination. Shifts in the composition and function of the microbiome are associated with many disease states including HIV, T1D, IBD as well as the response to therapeutic interventions such as anti-PD-1 cancer immunotherapy.

We leveraged a systems immunology approach on a 2018 cohort of YFV outbreak in Sao Paulo (Brazil) in order to better understand YFV disease severity and its’ potential causes. We combined plasma cytokine profiling, metabolomics and microbiome quantification to build a model of disease severity.
RESULTS

In order to better understand YFV-driven disease severity, we leveraged data from a cohort of YFV infected subjects from a Brazil outbreak in 2018. The cohort was composed of 87 survivors and 21 fatalities, aged between 18 and 87 (mean: 42.6; s.d.:15.14). It is composed of 85% males, which is a hallmark of silvatic YF due to more common outdoor working habits in rural areas. Lethality was found to be enriched slightly in older individuals (mean of 55 y.o. vs 42 y.o). Each patient self-reported the time of the onset of their symptoms upon their hospitalization. Given the debatable reliability of this self-reporting coupled with the heterogeneity for the timepoint at which they were hospitalized, we opted to normalize timepoints on Days since hospitalization (Suppl Fig. 1). We then proceeded to a clustering analysis on peak values of available clinical markers of multi-organ failure to detect survivors with a more severe disease presentation. We included International Normalized Ratio (INR), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (BT) and creatinine as markers of liver failure. The latter was also used as a surrogate of kidney failure. Moreover, platelet levels, erythrocytes, hematocrit and hemoglobin were used as markers of coagulopathy. We observed 2 clusters of subjects on the basis of these parameters, one of which contained all but 2 patients with a fatal outcome (Figure 1A). This cluster, which exhibited elevated levels of INR, ALT, AST, BT and creatinine, along with reduced coagulation markers, was consistent with that of a multi organ failure phenotype. Thus, we relabeled survivors contained within that cluster as survivor severe, while survivors in the other cluster we labelled as survivors mild.

We also explored established clinical scoring systems of organ failure in order to have 1) a more objective measurement of severity in relationship to the concentration values of these clinical markers instead of their relative abundance, but to also 2) allow the characterization of severity as a continuous variable instead of multiple subgroups of lower sample size. Subjects displaying
an elevated score for the Model for End-Stage Liver Disease (MELD), traditionally used in the context of chronic liver failure. It was also shown to be a good marker of acute liver injury in the context of SARS-CoV2 infection. In parallel, the Sepsis-Induced Coagulopathy (SIC) score, and the Estimated Glomerular Filtration Rate (EGFR), respectively markers of coagulopathy and kidney function, were also found to be elevated in patients in the severe cluster. We opted to use the MELD score as a primary outcome for our study of severity, given YFV’s hepatic tropism.

We hypothesized that YFV disease severity would be associated to a cytokine storm response, as observed in the context of SARS-CoV2. We thus quantified longitudinally the concentration of plasma cytokines and chemokines using the Mesoscale Discovery platform (Figure 1B). We observed 2 clusters of analyte expression, both of which were positively associated with severity. Multiple proinflammatory cytokines, including IFNa2a, IFN-b, IL-6, IL-18 and IGN-g were enriched in patients a lethal outcome or survivors with a severe presentation, and were correlated with the MELD score. We observed the highest levels of association of on the Day of hospitalization: we cannot discern whether that is because the effect is already at its peak early on followed by a resolution of inflammation, or is caused by the reduction in statistical power due to lower sample size as time passes due to the passing of patients and or failure to follow up.

We asked whether the drastic distinction between severe and mild cases could be ascribed to a difference in viral load (Figure 1C). We were surprised to observe no discernable association of viral load with severity at hospitalization. However, we observe a gradual increase in association to severity over time. This result suggests that differences in viral load do not account for the initial cytokine storm phenotype. However, it also suggests that the cytokine storm potentially impedes the resolution of viremia. Interestingly, a closer look at the per patient longitudinal
MELD score progression versus cytokine expression (Suppl. Fig.2) allowed us to observe that, for a subset of subjects, the cytokine storm precedes the multiorgan failure, which ultimately gives us insight into the chronology of disease progression.

Given the lack of association of viral load to severity at the initial timepoint, we attempted to find the cause for the discrepancy in inflammatory signatures across subjects. We hypothesized that some subjects were predisposed to a severe outcome due to environmental factors. One such factor is the microbiome, which has a well-established immunomodulatory function. Furthermore, the gut-liver axis, which corresponds to the bidirectional relationship between the liver and the microbiota, has been thoroughly studied. Furthermore, hepatic proinflammatory conditions, such as cirrhosis, have been known to perturb the gut microbiome and increases in microbial translocation. Thus, we quantified plasma markers of microbial translocation (Figure 2A) Intestinal fatty-acid binding protein (I-FABP), lipopolysaccharide binding protein (LBP) and soluble CD14. All three markers were significantly associated to the peak MELD score, indicating a heightened level of microbial translocation in severe disease.

In order to determine if the composition of bacteria differed across severity groups, we performed plasma Pathogen Sequencing (Pathseq) on a preliminary 17 samples, which allowed us to detect 183 bacterial taxons at the genus level. We quantified beta-diversity using the Bray-Curtis dissimilarity index, which allowed us to detect compositional distinctions in an unsupervised fashion in relation to disease severity, as depicted across the MDS1 axis (Fig.2B). We also quantified Shannon index as a measure of alpha diversity to assess whether samples obtained from severe disease were less diverse, as is often the case in pathological conditions such as IBD, T1D and even sepsis. We were able to show that, consistent with these conditions, lower diversity is significantly associated with a higher MELD score. Finally, a supervised analysis at the order level allowed us to negatively associate 2 types of proteolytic
bacteria, namely Micrococales and Rhizobiales, respectively of the Actinobacteria and Alphaproteobacteria phylae (Figure 2C). Collectively, these results indicate a strong relationship between YFV disease severity and the state of the microbial flora.

Given that the microbial immune interplay is often modulated through cell metabolism, we proceeded to quantify the plasma metabolome in YFV infected patients. We quantified 867 plasma metabolites using mass spectrometry, and performed unsupervised Principal Component Analysis (PCA) analysis (Figure 3A). We observed an evidence association of metabolic composition in relation to severity, with the MELD score clearly separating mild from severe cases along PC1 from left to right. We then proceeded to a supervised analysis followed by a pathway analysis using the MSEA approach to detect classes of metabolites enriched in severe or mild patients. We then used EnrichmentMap to collapse redundant individual significant pathways into 24 modules (Figure 3B). Of those 24 modules, 11 were enriched in mild disease whereas 13 were enriched in severe disease. We noted an interesting overrepresentation of amino acid metabolism pathways across significant modules. Indeed, aromatic amino acids tyrosine, tryptophan and phenylalanine were enriched in patients with higher MELD scores, whereas Branched-Chain amino acid (BCAA) were associated with a reduced severity. Moreover, additional key metabolic immunomodulatory pathways were also isolated, such as TCA cycle, and Acyl carnitine metabolism. Interestingly, phosphatidylcholine metabolism, which we identified in a currently under-preparation study of dengue vaccination as an inducer of ant-inflammatory responses through TGFβ signaling, was associated with milder disease. A closer look at the metabolites driving the enrichment of these key pathways (Supplementary Figure 3A) shows that citrate, phosphate, pyruvate and isocitrate, but not succinate, are associated to severe disease. This was a surprising finding, given succinate’s established role in the induction of IL-1B production. Overall, these findings support our
hypothesis that the microbial immunomodulatory function is regulated through metabolism in the context of YFV disease.

In order to build a mechanistic model of disease severity, we performed multi-Omic integrative analysis using sparse least-square regression modeling. We identified a clear association of the cytokine storm response (cluster 1, MSD) with histidine, tyrosine, tryptophan and lysine metabolism, and a negative association with phosphatidylcholine and long chain fatty acid metabolism. We also observed a significant association of Micrococal abundance in the plasma to this cytokine storm signature, supporting a functional interplay between the immune component, the metabolome and the microbiome in the context of severe disease. Moreover, we observed an association of immature neutrophil counts, which were previously identified as markers of severity in [REF], with both the cytokine storm and the Micrococal.

DISCUSSION
In this study, we leverage a systems immunology multi-Omic approach to explore the mechanistic determinants associated with YFV Hemorrhagic Fever severity. We used a cohort of YFV infected individuals in Brazil to identify a substantial cytokine storm response mounted in a subset of individuals with multiple signs of multi-organ failure. We leveraged unsupervised analyses to classify patients into severity groups in combination with different established clinical scoring systems of organ damage, such as the MELD score. Prior approaches compared fatal outcomes with survivors to identify pathological determinants, which limits the sensitivity of detection due to the heterogeneity across survivors, which range from very mild to very severe. Thus, using scoring systems like the MELD allows to better predict the severity of disease, and identify associated factors with increased sensitivity.
It has been hypothesized for a long time that YFV infection can lead to a cytokine storm response: we have demonstrated it for the first time in the context of in vivo infection. We detect early on a massive presence of proinflammatory cytokines in the plasma in the context of severe disease, which was not seemingly caused by a difference in viral load. Despite viral load being a predictor of fatality in previously published work, this appears to be caused by the cytokine storm as discrepancies in viremia follow chronologically the onset of the cytokine syndrome.

We have identified an increase in gut permeability and microbial translocation in the context of severe disease, which was captured by multiple markers. This reflects observations from chronic hepatic inflammation, such as alcoholic cirrhosis and non-alcoholic steatohepatitis. Moreover, there are reports of increased LPS levels in the context of hepatic viral infections, such as Hepatitis B and C. The increased endotoxins could therefore contribute to the systemic proinflammatory response. Moreover, the microbial composition of the translocated components also seem to play a role in disease pathology: not only do more bacteria translocate, but we notice a shift in which bacteria do. The negative association of 2 proteolytic bacteria orders with YFV severity, namely Rhizobiales and Micrococales, highlights how distinctions between individuals can bring light as to how some individuals mount a cytokine storm response whereas others do not. Proteolytic bacteria are known for their capacity to use exogenous amino acids as a source of energy, such as the hosts.

It thus becomes intriguing to observe a number of amino acid metabolism pathways being associated with disease. Indeed, we observed an enrichment of tryptophan, tyrosine and phenylalanine related metabolites in severe outcome, in contrast with an increase in branched chain amino acids like leucine, isoleucine and valine in milder disease. How that translates into modulation of a proinflammatory response is not clear yet, although we can provide an
hypothesis. The metabolic requirements, both in terms of ATP and amino acid supply, for
mounting a large scale immune response as is observed in a cytokine storm are substantial.
Therefore, bacteria known to contribute to the modulation of that metabolism would potential
compete for resources with the immune system. In that context, we hypothesize that proteolytic
bacteria effectively starve out the cytokine storm response in milder disease. However, we do
not yet understand the differential impact of distinct classes of amino acids in this process,
which will be the subject of future studies. We also observe an association of the TCA cycle with
severe disease: we cannot yet dissect if this is caused by the added ATP being made available
downstream of this process, or by the immunomodulatory nature of some known TCA cycle
intermediates (such as succinate). Moreover, we have to consider that TCA cycle and amino
acid metabolism are closely intertwined biologically, given that the latter can be used to
generate TCA cycle intermediates after deamination.

We propose a model where individuals with undistinguishable initial viremia initiate a
proinflammatory response similar in magnitude. However, severe outcome begins to differ from
a mild presentation due to interindividual variations in neutrophil numbers, leading to higher
downstream activation of immune activation. In both cases, this leads to an increased gut
permeability, which favors the translocation of certain bacterial species that can modulate the
metabolic supply of amino acids and ATP required for a full scale proinflammatory response. As
such, a milder presentation is limited by the lack of resources. The cytokine storm response is
known to negatively impact the CD8 T cells response: this would further influence how viremia
is controlled over time, leading to a higher viral load at later timepoints in severe disease,
favoring a forward loop of microbial translocation and metabolic modulation of the immune
system.

FIGURE LEGENDS
**Figure 1. Cytokine Storm is associated with YFV Disease severity.** A) Heatmap of z-score normalized clinical markers of YFV disease and multi-organ failure. Each row represents a separate marker, and rows individual subjects at the peak of their response. A blue color represents a lower expression of each marker, while red denotes a higher-than-average level of detection. Column annotations (above) display independent peak clinical scores per patient (EGFR, MELD and SIC), or Severity Subgroups inferred by clustering analysis. B) Longitudinal plasma analyte expression heatmap in YFV infected subjects. Each row represents a z-score normalized analyte, and each column represents a sample. Blocks of samples are separated by timepoint, and column annotations as described in A. Cluster membership for rows was inferred by the Gap statistic and kNN clustering. C) YFV viral load scatter plots per timepoint (Days since hospitalization) versus MELD scores. Each point represents a patient, with the color showing their respective maximum Viral load across timepoints.

**Figure 2. Gut-liver axis breakdown leads to increased microbial translocation and differential microbial composition in the context of severe YFV disease.** A) Scatter plot of 3 plasma markers of microbial translocation versus the peak MELD score. B) Pathseq beta-diversity (Bray-Curtis) of plasma Pathseq data from 17 YFV infected patients. Each dot represents a subject at or close (+2 days) to hospitalization timepoint. Color maps to the peak MELD score per subject. C) Pathseq Shannon alpha diversity scatter plot versus the MELD score. Each dot represents a patient: color is mapped to fatality subgroups. D) Scatter plot of microbial abundance for 2 orders of significantly associated bacteria versus the peak MELD score. Dots represent subjects and are colored as described in C.

**Figure 3. Metabolic activity is associated with YF disease severity.** A) Principal component analysis of plasma metabolite detection in YFV-infected subjects. Each dot represents a patient, and the x and y axes represent PC1 and PC2, respectively. Dot color is mapped to the MELD
score, while dot size is mapped to patient age. Dot shape is mapped to patient fatality subgroup.

B) Heatmap of plasma metabolite pathway analysis. Each row is a metabolic module built with the combination of MSEA and EnrichmentMap, while columns represent individual subjects. Colors represent averaged row-normalized expression of metabolites per module, from blue to red representing low to high expression, respectively. Column annotations represent Fatality subgroups and MELD score.

**Figure 4. Integrative analysis builds an interconnected mechanistic model of YFV disease severity.** MultiOmic inter-feature association is represented by a Circos plot of significantly correlated features. Each sector represents a feature significantly associated to the MELD score (outer circle color), with their respective source OMIC depicted in the inner circle. Links between sectors represent a significant association, with the colors mapped to the signed(-log(p')). As such, a blue link or a red link respectively depict a negative or positive association between features.

**Supplementary Figure 1**: Distribution of YFV infected cohort across timepoints, both on days since onset of symptoms (left) and Days since hospitalization (right). Subjects with a fatal outcome are displayed in pink while survivors are in turquoise.

**Supplementary Figure 2**: Longitudinal scatter plot of averaged cluster expression per subset versus the MELD score. Color is mapped to the plasma analyte levels per condition as a gradient, with blue denoting a lower expression and red a high expression, respectively. Lines link subjects across timepoints.

**Supplementary Figure 3**: Heatmap of selected leading edge plasma metabolite identified in pathway analysis. Each row is a plasma metabolite, while columns represent individual
subjects. Colors represent row-normalized expression of metabolites, from blue to red representing low to high expression, respectively. Column annotations represent Fatality subgroups and MELD score. Row annotations represent from which module each metabolite was enriched in.
Supplementary Figure 1.
Figure 1.
Supplementary Figure 2.
Figure 2.
Figure 3.
Supplementary Figure 2.
Figure 4.