Can we make learned decisions and find alternative cure options for the novel coronavirus disease?

COVID-19 is a severe acute respiratory syndrome (SARS) viral disease caused by SARS-CoV-2 virus. Like many other SARS coronaviruses, the SARS-CoV-2 attacks the lungs and may eventually lead to either direct viral pneumonia or secondary bacterial pneumonia. In the worst-case scenario, when lungs are damaged by the SARS-CoV-2, critically ill patients require artificial ventilation to deliver enough oxygen to the body. And, here comes the first and difficult question: in case of limited resources, how to decide which patient should be attended to, and which patient is to reject?

Objectively this can be decided by the development of machine-learned diagnostic models, but at the moment we don't have many available data from COVID-19 patients. However, we can bypass this by taking existing data available for bacterial pneumonia patients, which might provide a good approximation to the SARS-CoV-2 induced pneumonia. As such, we used externally available blood genome-wide gene expression profiles collected for 183 pneumonia patients treated in the intensive care unit[1]. For 40 patients (22%) treatment had failed and they died within 30 days. Thus, we wondered if we could predict from the blood gene expression profiles collected during the treatment who of the bacterial pneumonia patients would have better chances of the survival?

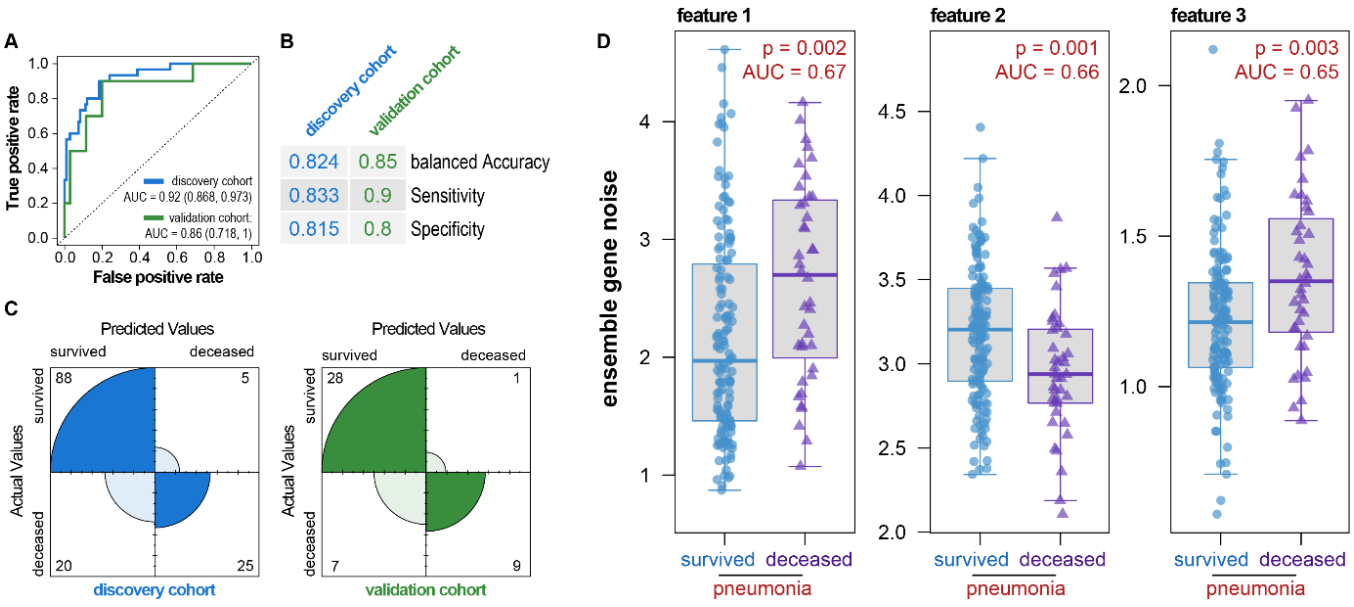
First, we noted that age of survived and deceased patients did not differ significantly ($p = 0.22$). This means that the age of the patient *per se* cannot be valid criteria to predict pneumonia survivorship. Thus, we turned into analysis of blood gene expression profiles represented by RNA copy numbers for each gene of every patient. We reasoned that the critical patient's state could be learned from imbalances in RNA copy numbers, and here we take the word "imbalance" literally. We assume that like in Swiss watch, gene ensembles should express coherently. For example, genes encoding subunits of some protein complex or proteins of some biological pathway should express

[1] B.P. Scicluna *et al*., American Journal of Respiratory and Critical Care Medicine 192 (7), 2015

in specific proportions, *i.e.* stoichiometries. However, gene expression is a stochastic process, which results in spontaneous deviations in RNA copy number or gene noise, and this may be responsible for critical imbalances in gene ensembles expression[2]. Thus, unlike a traditional approach, where specific changes in gene expression are evaluated (differential gene expression analysis), we treat this problem from the view point of gene ensembles and their gene noise.

Estimating gene noise for known gene ensembles (genes with similar biological or biochemical function) we trained a diagnostic model, which predicts the survival and mortality of bacterial pneumonia patients with high accuracy (balanced accuracy > 80%). To this end, we randomly split patients into discovery (138 patients, 75%) and validation cohorts (45 patients, 25%). The gradient boosting classification machine has been trained on the discovery cohort using gene noise for the selected gene ensembles as feature space, and then independently tested on validation cohort.

Figure 1



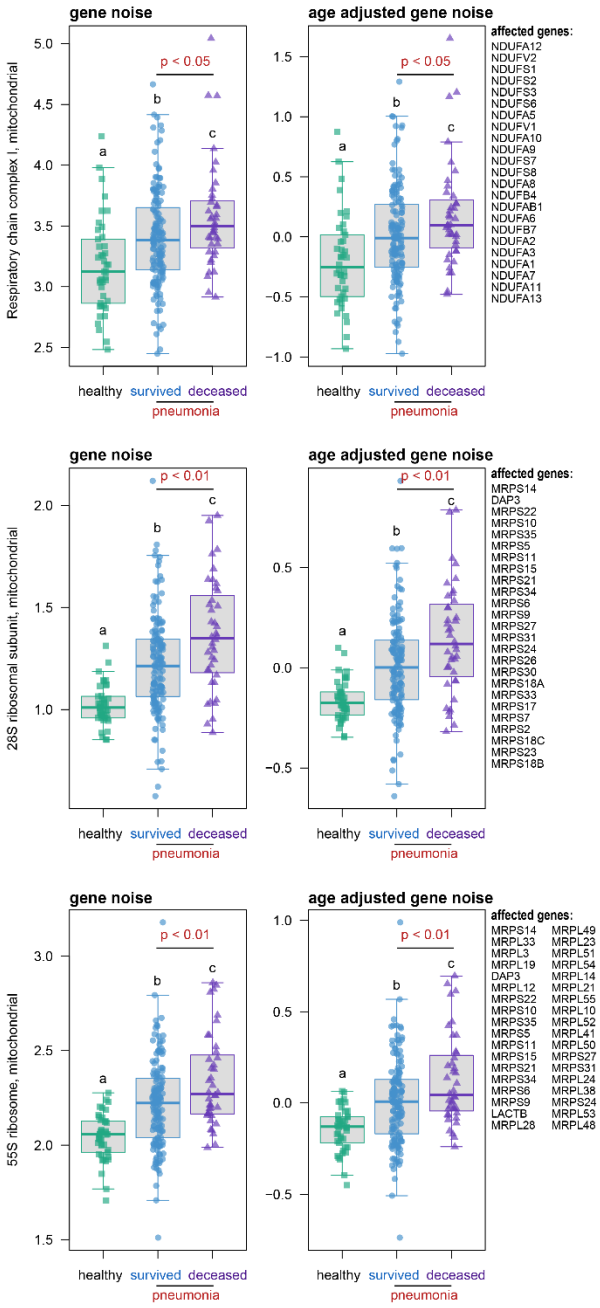
From Figure 1 it is clear that our gene noise model is able to predict mortality outcome for bacterial pneumonia patients with acceptable for diagnostics accuracy. For model accuracy, we used conventional measures such as area under the receiver operating characteristic curves (AUCs), Specificity, Sensitivity and balanced accuracy (bACC = (Specificity + Sensitivity)/2) for both discovery and validation cohorts (Figure 1A, B). Confusion plots illustrating true positive/negative

[2] T.V. de Jong, Y.M. Moshkin, V. Guryev., Physiological genomics 51 (5), 2019

and false positive/negative cases are shown in Figure 1C. In diagnostics, high specificity (true negative rate) implies that the model is able to discriminate correctly negative cases or, in this case, potential survivals. High sensitivity indicates that the model identifies correctly positive cases, or, in this case, patients with high probability not to survive. It also has to be noted that even individual ensemble gene noise features are capable of discriminating patients with high and low chances of survival with reasonable accuracy (Figure 1D).

These individual features can be used to develop faster and cheaper PCR diagnostic assays as compared to more expensive and slower genome-wide assays. Thus, we conclude that, although more studies are needed, in principle it is possible to predict mortality outcome for either coronavirus (COVID-19) or bacterial pneumonia patients.

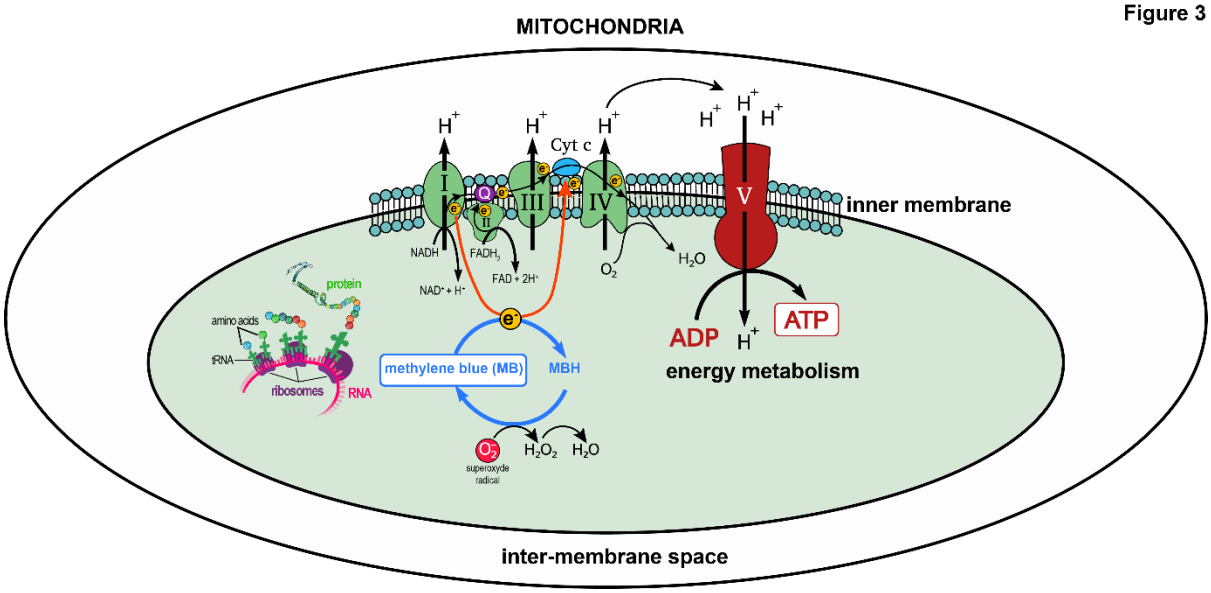
Figure 2



But what else can we learn from the gene noise? Apparently, we can discover for new targets for cure. Traditionally pharmacological sciences are focused on specific druggable targets, *i.e.* specific proteins and enzymes, which either mutated or mis-regulated. However, an idea of gene ensemble gene noise allows to look at the disease under completely different angle. Indeed, gene noise is totally independent of the mean RNA copy number (a value usually estimated in biology and pharmacology), and gene noise treats a pathology as a result of imbalanced expression of ensembles of genes rather than individual genes. In other words, while mean expression of genes in ensemble may not be even affected by a disease, their ensemble gene noise may well be.

Considering pneumonia patients, we noted that ensemble gene noise has been increased significantly for genes encoding mitochondrial respiratory complex I and mitochondrial ribosome 28S and 55S complexes as compared to healthy individuals (Figure 2 and 3). Because healthy individuals were on average younger than pneumonia patients, we also calculated age adjusted ensemble gene noise (Figure 2, right panel). Thus, we conclude that pneumonia results in imbalanced expression in genes encoding for key mitochondrial proteins. Moreover, in patients with low survivorship (deceased individuals within 30 days of treatment) ensemble gene noise for these mitochondrial complexes (respiratory complex I, and 28S and 55S ribosome subunits) was further increased. This suggests that compromised balance in expression of these genes contributes to a) pneumonia pathology and b) survivorship of pneumonia patients.

Figure 3



Could these genes be potential druggable targets for pneumonia patients including severe COVID-19 patients? We believe yes. As such, one could pay attention to drugs which improves mitochondrial respiration. One of such off-shelf drugs is Methylene Blue (MB). MB acts as a redox-active alternative electron donor/acceptor that bypasses respiratory complex I/II in electron transfer chain (Figure 3). In other words, if the function of mitochondrial respiratory complex I is compromised due to elevated ensemble gene noise in severe pneumonia patients, this can be alleviated by MB. In fact, studies on model animals (mouse, rat and others) seem to support this by

showing that MB might be beneficial in treatment of aspiration pneumonia[3]. Of course, MB won't treat viral infection itself, but due to its ability to improve mitochondrial respiration it might act as potential adjuvant to lessen severity of pneumonia symptoms.

In line with this, other potentially druggable pathways might be considered from the analysis of ensemble gene noise. The additional data will be available upon request.

Sincerely yours,

Yuri Moshkin, CSO of Gene Learning Association
yury.moshkin@gmail.com

[3] M. Kanter *et al*., Thoracic 193(2), 2015; O.V. Evgenov *et al*., Eur. Respir. J. 20, 2002