Melatonin-Index as a biomarker for predicting the distribution of presymptomatic and asymptomatic SARS-CoV-2 carriers

Pedro A. Fernandes1*; Gabriela S. Kinker2; Bruno V. Navarro3; Vinicius C. Jardim4, Edson D. Ribeiro-Paz1, Marlina O. Córdoba-Moreno1, Débora Santos-Silva1, Sandra M. Muxel1, Andre Fujita4, Helder I. Nakaya5, Marcos S. Buckeridge3, Regina P. Markus*1

1 Laboratory of Chronopharmacology, Department of Physiology, Institute of Bioscience, University of São Paulo, Brazil
2 International Research Center, A.C. Camargo Cancer Center, São Paulo, Brazil.
3 Laboratory of Plant Physiological Ecology (LAFIECO), Department of Botany, Institute of Biosciences, University of São Paulo, São Paulo, Brazil.
4 Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, São Paulo, Brazil.
5 Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil
*Correspondence: pacmf@usp.br; rpmarkus@usp.br, Tel: +55-11-30816996

Running Title: Lung melatonin modulates SARS-CoV-2 infection

Received: August 6, 2020; Accepted: January 2, 2021

ABSTRACT

The pandemic dissemination of the SARS-CoV-2 led, on the one hand, to a worldwide effort to develop mechanistic-based therapeutics and vaccines, and on the other hand, the searching for determining the spreaders and the mechanisms of transmission. Melatonin, a multitask molecule, orchestrates defense responses by allowing the proper mounting, duration, and magnitude of innate immune responses. Melatonin is synthesized on demand by immune-competent cells and constitutively by resident macrophages such as alveolar macrophages. Here we investigated whether the expression of genes relevant to virus invasion and infection varies according to a genic index (MEL-Index) that estimates the capacity of the lung to synthesize melatonin. A COVID-19-Signature composed of 455 genes of 288 human lungs (GTEX, UCSD) was correlated with MEL-Index by Pearson correlation test, gene-set enrichment analysis, and networking tool that integrates the connectivity between the most expressed genes, allowing us to compare the same set of genes under different states. The three independent procedures point to a negative relationship between MEL-Index and SARS-CoV-2 infection. The entry in epithelial AT2 cells should be hampered by a positive correlation TMRPSS2 and a negative correlation with the coding gene for furin, suggesting dysfunctional processing in virus spike. Moreover, MEL-Index also negatively correlates with the genes that codify the proteins of multi-molecular receptor complex CD147, the gateway in macrophages, and other immune cells. In summary, the perspective that lung and respiratory tract melatonin could be a natural protective factor opens new epidemiological and pharmacological perspectives, as high MEL-Index scores could be predictive of asymptomatic carriers, and nasal administered melatonin could prevent evolution of presymptomatic carriers.

Key Words: SARS-CoV-2 transmission; lung melatonin; MEL-Index score; virus infection.
1. INTRODUCTION

The pandemic dissemination of the SARS-CoV-2 led to a worldwide effort to develop mechanistic-based therapeutics and vaccines (1). Virus gateways, proliferation/dissemination pathways, and host targets responsible for dysfunctional defense responses were explored as possible targets for treating COVID-19 patients. The repurposing of drugs targeting the virus is still a promise, and the most widely approved strategies are based on supportive treatments such as anticoagulants (2), antagonists of the IL-6 receptor (tocilizumab), (3) and dexamethasone in critical patients (4). Vaccines are available, but SARS-CoV-2 mutations (5) increased the possibility of transient immunity and vaccine scape. Despite all the efforts for combating the disease, the second pandemic wave at the end of 2020 urgently demands new approaches for understanding virus transmissibility, infection, and replication.

Epidemiologic studies confirm that asymptomatic and presymptomatic carriers are hidden spreaders (6). The percentage of asymptomatic and presymptomatic spreaders in small and isolated populations present a variation not well understood. Analysis of virus RNA load by PCR in two outbreaks in a Japanese and an Argentinian cruise ship confirmed the presence of asymptomatic, presymptomatic, and symptomatic COVID-19 patients. (6). The percentage of asymptomatic and presymptomatic individuals varied from 30 – 80%, depending on the cruise (7,8), strongly reinforcing the hypothesis that not yet identified mechanisms may coordinate virus airborne invasion and infection on an individual basis.

Airborne SARS-CoV-2 transmission was dichotomized in short-distance (larger droplets) and long-distance (aerosol), and productive invasion of the airways relies on detection of virus genetic material, while infection activates innate and acquired immune responses, resulting in the production of antibodies (9). Despite the great number of systematization of clinical case reports, the repurposing of drugs, and the determination of the cellular and systemic mechanisms responsible for the installation and degree of severity of the disease, the pandemic is not controlled, many countries imposed new isolation during the last days of 2020. It is evident that new biomarkers need to be established in order to detect hidden or camouflaged spreaders.

Melatonin (N-acetyl-5-methoxytryptamine) is involved both in the mounting and the recovery of innate immune responses. Just at the beginning of the process, pineal melatonin synthesis is blocked, and neutrophils migrate from the bone marrow to the circulation and from the circulation to the injured area (10, 11). Mitochondria phenotypes and functionality are regulated by cell-synthesized or circulating melatonin (12), which orchestrates the balance between efficient energy production and building material for repair, growth, and division. Virus infection shifts this balance to aerobic glycolysis, ensuring building material for replication. This mechanistic evidence plus converging epidemiological profiles of groups with reduced melatonin production and high COVID-19 risk (13) subsidized the proposal to use melatonin as an adjuvant in the treatment of COVID-19 (14). Clinical trials for evaluating the protective effect of melatonin as the darkness hormone (15), and supporting drug for treating severe cases of COVID-19 (16).

The previous studies with melatonin did not consider lung synthesized melatonin as a relevant source for allowing the first stages of invasion and infection, opening the possibility that lung melatonin could be relevant for defining the distribution of presymptomatic and asymptomatic carriers. Considering that melatonin synthesized by rodent alveolar macrophages exposed to pathogen- and damage-associated molecular patterns (PAMPs, DAMPs) impairs their activation (17). A two-gene index score (MEL-Index) that predicts the severity of human glioma (18, 19) was used as a biomarker for estimating the expression of genes related to airborne SARS-CoV-2 invasion and infection in normal human lungs RNA-seq data deposited in the Genotype-Tissue Expression (GTEx). Our data provide the first
biological basis for considering lung melatonin synthesis a positive prognostic marker for predicting the distribution of asymptomatic and presymptomatic carriers.

2. MATERIALS AND METHODS

2.1. Genes analyzed.

A total of 455 genes shown to be associated with COVID-19 comorbidities (20), virus/human proteins interaction (21), virus gateway (22, 23) were named “COVID-19-Signature”. A COVID-19 physiological signature was defined by 212 genes involved in the entry of SARS-CoV-2 in human cells (54), intracellular traffic (66), mitochondrial activity (42), and genes linked to transcription and post-translation processes (79).

2.2. RNA-seq data.

Gene-expressed in healthy human lungs were obtained from the Genotype-Tissue Expression (GTEx) RNA-seq data, downloaded from the UCSC RNA-seq Compendium (10.1038/nbt.3772), deposited at the UCSC Xena Browser (https://doi.org/10.1038/nbt.3772). Expression levels were quantified using RSEM, upper quartile normalized, and log2 (fpkm+0.001) transformed. In total, we analyzed 288 samples of healthy GTEx lung samples.

2.3. MEL-Index calculation.

MEL-Index is based on the relative expression of genes that codify the enzymes to convert N-acetylserotonin to melatonin (ASMT, acetyl-serotonin methyltransferase) and melatonin to 6-hydroxymelatonin (cytochrome P450 (CYP)1B1) in brain, liver, and other tissues (18, 19, 24). The ASMT and CYP1B1 levels were used by z-normalizing the log2 transformed values of the genes and calculating the difference between them (ASMT - CYP1B1) for each sample.

2.4. Animals & Lung melatonin content.

Adult male Wistar rats (2–3 months old, 200–250 g) maintained in 12/12 light-dark cycle (lights on at 06h00) received water and food at libitum. The rats were killed at 9h00 and 12h00 by decapitation at daytime when no pineal melatonin contribution is expected. Lungs were immediately frozen in liquid nitrogen (-170°C) and then maintained at -80°C till evaluation. The procedures followed the ethical standards of the Brazilian National Council on Experimental Animal Control (CONCEA) and were approved by the local Ethics Committee (protocol 346/2019). Frozen lungs homogenized in Tris-HCl buffer (pH 7.4, 1mM EGTA, 1mM EDTA) were centrifuged (20,000g, 5 min, 4°C). Melatonin and proteins were quantified in the supernatant with saliva ELISA kit (IBL, Hamburg, Germany, REF: RE54041), and protein Bradford assay (Bio-Rad, California USA, REF: #500-0006).

2.5. Real-time polymerase chain reaction (qPCR): Relative quantification.

Total RNA was extracted from 50 mg of each lung homogenized in liquid nitrogen, according to the TRIzol® (Invitrogen) manufactory’s instructions. Two µg of RNA samples were reverse-transcribed to form complementary DNAs by the SuperScript III First-Strand Synthesis System (Invitrogen, USA) and one ng of random primer (Invitrogen, USA), following the manufacturer’s instructions. The qPCR was performed using 2X SYBR Green PCR Master Mix (Applied Biosystems), 200 nM of each primer pair, 5 μL of cDNA (10-fold
diluted), and RNase-free water to a final volume of 25 μL. The reactions were performed in StepOne Plus Real-Time PCR System (Applied Biosystems, USA). The mixture was incubated at 94°C for 5 min followed by 40 cycles of 94°C for 30 s and 60°C for 30 s. The following primer pairs were used for rat mRNA analysis: Asmt: 5’-agcgcctgtctgtcatgag-3’; 5’-ggaagcgtgagaggtcaaagg-3’, Cyp1b1: 5’-gacatcttgagccagcca-3’; 5’-aaagccatgacgtatgtgtaagtt-3’ and Gapdh: 5’-gcccagaacatcatccctg-3’; 5’-ggaacacggaaggccatg-3’. The 2(-ΔΔCt) formula was applied to calculate the relative expression of Asmt and Cyp1b1 normalized by Gapdh. The mean ΔCt of the control group was used as the calibrator.

2.6. Analytical methods and statistical procedures.

2.6.1. Correlation between the COVID-19-Signature score and the MEL-index score.

In each dataset analyzed, genes were ordered according to the mean expression across samples and divided into 50 bins. Next, we generated a control index score, for each sample, by subtracting the average z-normalized expression of 100 genes randomly sampled from the CYP1B1 bin from the average z-normalized expression of 100 genes randomly sampled from the ASMT bin. Finally, for each sample we created a COVID-19-Signature score, defined as the average z-normalized expression of the selected genes. We then calculated, for each dataset, the correlation between the COVID-19-Signature score and the MEL-index score, as well the control index score. The same statistical procedure was applied to subgroups of COVID-19-Signature related to virus interaction and cellular processes linked to virus invasion. Statistical analyses were performed with the GraphPad Prism 6 or the R software (http://www.r-project.org). To determine correlations, we calculated Pearson’s correlation coefficient. P values < 0.05 were considered statistically significant.

2.6.2. Gene set enrichment analysis.

GSEA was performed using the desktop application v4.0.3 (25). Genes from lung samples were ranked according to Pearson’s correlation coefficient between their expression and the MEL-Index calculated for each sample. Enrichment scores (ES) were normalized (NES) to account for gene set size, and p-values were calculated using 1,000 gene set permutations.

2.6.3. Network analysis.

The MEL-Index allows the sorting of the samples according to the values obtained for each sample. We divided the 288 samples into Low (-3.904 to -0.109), Medium (-0.108 to 0.827), and High (0.828 to 3.356) MEL-Index groups and performed networks analysis for each subgroup (96 samples per group). Weighted networks were generated based on Pearson’s correlation with a cut-off of 0.001 p-values adjusted by multiple tests (False Discovery Rate - FDR). The nodes represented the genes selected, and the correlation levels represented the links. This analysis was carried out by BioNetStat package (26), and the complex network parameters degree centrality and degree distribution were evaluated.

3. RESULTS

3.1. Relationship between MEL-Index and lung melatonin content.

To validate that the lung genetic MEL-Index (Figure 1A, 13) could be a predictor of melatonin content, we determined the expression of the genes Asmt and Cyp1b1 and the
concentration of melatonin in the lung of rats killed during daytime (Figure 1B). The increase in melatonin levels in the lung observed between 09h00 and 12h00 was dependent on the increase in Asmt transcription, and was parallel to the increase of the MEL-Index (Figure 1B).

Our next step was determining MEL-Index/ COVID-19-Signature score correlation in 288 samples of healthy human lungs. MEL-Index significantly correlated with the COVID-19-Signature score \( (r = -0.15, p = 0.01) \), but not with the control index score \( (r = -0.03, p = 0.63) \). The percentage of negatively correlated genes with the MEL-Index (35.8%) was the double of the positively correlated ones (47% had no correlation), strongly suggesting that melatonin synthesized in the lung could mediate virus-induced response (Figure 1C). Functional enrichment analysis, commonly referred as gene set enrichment analysis (GSEA) is an alternative manner to statistically test the expression profile of a set of genes in function of a specific condition/variable. This analysis confirmed that MEL-Index negatively correlates with the expression of genes assigned as COVID-19-Signature \( (\text{NES} = -2.15, \text{qFDR} < 0.0001) \) (Figure 1C, Appendices Table S1).

![Fig. 1. MEL-Index score & human COVID-19-signature.](image)

(A) MEL-Index formula. (B) MEL-Index as a predictor of rat lung melatonin concentration. Asmt (left green) and Cyp1b1(red) mRNA expression, Asmt minus Cyp1b1 (MEL-Index) combined expression (blue) and melatonin content (right green) of rat lungs. Data are expressed as mean ± sem (\( N = 3 \) per point). Data are expressed as mean ± sem. (C) MEL-Index negatively correlates with COVID-19-Signature and functional enriched scores (288 human lung samples, MEL-Index score: -3.904 to 3.356).

TRP: tryptophan; TPH: tryptophan hidroxilase; 5-HTP: 5-hidroxytryptophan; AADC: aromatic L-amino acid decarboxylase; 5-HT: serotonin; AANAT: aryalkyl-N-acetyltransferase; NAS: N-acetylseryotonin, ASMT: acetylserotonin methyltransferase;
MEL: melatonin; CYP1A1: cytochrome P4501A1; CYP1A2: cytochrome P450 1A2 and CYP1B1: cytochrome P450 1B. ASMTi = individual ASMT expression; CYP1B1i = individual CYP1B1 expression.

3.2. MEL-Index and networks analysis.

The structures of networks generated with the COVID-19-Signature genes were evaluated in three MEL-Index subgroups (High, Medium, and Low MEL-Index). Degrees of centrality and distribution were analyzed in these groups using BioNetStat Bioconductor package (26).

The subgroups differed in the degree of centrality and distribution (Figure 2A, Table 1), with Low, Medium, and High MEL-Index groups presenting 25192, 43971, and 17681 links, respectively. Pairwise comparisons for degree distribution of the samples showed that the Medium subgroup differed from the others (Table 1). The low subgroup showed a higher density around degree 50, the Medium subgroup at degree 150, and the High subgroup at 25 links (Figure 2A). Moreover, besides reducing the connectivity per node, there was an increase in the number and percentage of negative links in the High MEL-Index subgroup (Figure 2B).

**Correlation Gene Networks According to Mel-Index**

![Network Analysis](image)

**Fig. 2.** Degree of expression and connectivity according to complex network analysis (BioNetStat, 26).

(A) Histograms and degree distributions of COVID-19-Signature genes ranked by Low (-3.904 to -0.109, red), Medium (-0.108 to 0.827, blue), and High (0.828 to 3.356, green) MEL-Index scores. (C) The number of positive and negative links observed in Low, Medium, and High MEL-index subgroups of normal lung tissues.
Table 1. Network analysis: Comparisons of the degrees of centrality and distribution in networks formed by the COVID-19-Signature genes.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Test</th>
<th>p-value</th>
<th>Q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low x Medium x High (Normal)</td>
<td>35.08</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>Low x Medium (Normal)</td>
<td>30.70</td>
<td>0.005</td>
<td>0.008</td>
</tr>
<tr>
<td>Low x High (Normal)</td>
<td>17.61</td>
<td>0.144</td>
<td>0.144</td>
</tr>
<tr>
<td>Medium x High (Normal)</td>
<td>44.83</td>
<td>0.001</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Highlighting the interaction of the top 40 most connected genes (Appendices Table 2) in Low, Medium, and High MEL-index samples, we observed different profiles of interactions among the networks. The higher density of connections per node was observed in the Medium Mel-Index subgroup (Figure 3A, larger nodes), while in the High MEL-Index subgroup, the number of negative links (red lines) was more evident. Moreover, the genes presenting the highest number of significant connections differ among the groups, with only three genes appearing in all the subgroups (Figure 3B).

![Network Analysis](image)

**Fig. 3.** Visual image of the top 40 most connected genes of Low, Medium, and High MEL-Index of NORMAL human lung.

(A) Graphs showing the connections of the subgroups presenting the strength of each node with the genes of the COVID-19 signature. The genes positions are conserved among networks to allow the comparison among them. (B) Venn Diagram showing the number of nodes shared among the top 40 most connected genes of Low (-3.904 to -0.109), Medium (-0.108 to 0.827) and High (0.828 to 3.356) MEL-Index subgroups. The size of the nodes reflects the density of connections; blue lines represent positive and red lines negative interactions.
3.3. MEL-Index and functional groups.

Our study's main goal was to determine if melatonin production in the lung, estimated by the MEL-Index score, could confer previous protection against the invasion of epithelial cells and alveolar macrophages. Therefore, the influence of MEL-Index on the expression of genes involved in SARS-CoV-2 virus entry in lung epithelial AT2 cells and alveolar macrophages (virus gateways, 54 genes), together with those relevant for intracellular trafficking (66 genes), transcription/(post)translation (79 genes) and mitochondrial function (42 genes), selected among the COVID-19-Signature genes (Appendices), were analyzed as a subgroup.

Most genes of the physiological signature were negatively correlated with the MEL-Index (46.7%, multi-correlation test, P<0.05) (Figure 4A). When both the error of rejecting a correct hypothesis and that of accepting a false hypothesis (FDR<0.1) were considered, 25% of the genes negatively correlated with MEL-Index. The percentage of negatively correlated genes was much lower (17%) than that of neutral genes (36.3%). GSEA confirmed the negative correlation between all the genes related to the physiological functions, strongly suggesting that high lung melatonin synthesis will result in higher protection against virus (Figure 4B).

![Figure 4. MEL-Index & Physiological signature.](image)

(A) Number and percentage of positive, neutral and negative correlations with the MEL-Index. (B) GSEA analysis of 212 genes composing the Physiological signature gene set. (C) Number and percentage of positive, neutral and negative correlations with the MEL-Index in the Virus gateways, Intracellular traffic, Transcription/(post)Translation and Mitochondrion subgroups. Gene numbers within parentheses.

The next step was to discriminating the genes inside each group. Overall, 33.4%, 53.1%, 50.7%, and 57.1% of the genes associated with virus gateway, intracellular traffic, transcription/(post)translation regulation, and mitochondrial physiological functions,
respectively, were negatively correlated with the MEL-index according to the probability of rejecting the nullity hypothesis in a multi-correlation test (P<0.05). When both the error of rejecting a correct hypothesis and that of accepting a false hypothesis are considered (FDR<0.1), the percentage of negatively correlated genes with the MEL-Index are still around 30% in the intracellular traffic, transcription/(post)translation regulation, and mitochondrial physiological functional groups and of 16.7% in the virus gateway group. Accordingly, the percentage of neutral correlations varied from 30 to 45 % in the groups, and the percentage of negative correlations with the MEL-Index were very low (12.2%, 12%, 19%, and 7.2%) in the virus gateway (12.2%), intracellular traffic (12%), transcription/(post)translation (19%), and mitochondrion (7.2%) functional subgroups (Figure 4C). Of note, considering the SARS-CoV-2 receptors pathways, we observed that eight out of 11 genes associated with the CD147/CD44 membrane complex were negatively correlated with the MEL-Index (Appendices, Table S3). On the other hand, TMRSS2, associated with the ACE2 virus gateway, was positively correlated with the MEL-Index (Appendices, Table S3). As discussed below, these discrepant results might collaborate for the same protective effect, reducing the virus entry in human lung cells.

4. DISCUSSION

This year, governed by the SARS-CoV-2 pandemic, there were advances in understanding the virus entry in the lung via the AT2 epithelial cells, the immunological reactions according to the disease progression, and the development of vaccines. However, there is an important gap regarding the transmission between humans. What are the biological processes that define symptomatic and asymptomatic carriers? Is there any difference between asymptomatic carriers and those exposed to the same viral load that did not test positive for PCRs? The present study aimed to highlight differences in healthy human lung gene expression that could support preventive treatments to hamper invasibility and propose a biomarker for detecting asymptomatic carriers.

Alveolar macrophages that compose up to 90% of total macrophages in healthy human lungs lead to immune suppression and mucous production after phagocyting airborne microorganisms and particulate material (28). The phagocyted material is expelled together with mucous, protecting the individual against infection and the mounting of immune defense response. Melatonin synthesized in the respiratory tract by resident or monocyte-derived macrophages, besides impairing the mounting of innate immune responses against particulate material of polluted air and microorganisms, potentiates phagocytosis (17, 29-31). Considering that the MEL-Index score in healthy rats was directly related to melatonin content, we decided to explore whether the MEL-Index score from healthy human lung could be implicated in the biology of asymptomatic carriers.

We established a COVID-19-Signature composed of 455 genes of 288 samples of normal human lungs from the UCSC Genotype-Tissue Expression RNA-seq according to their relevance for aggressiveness and severity of the disease described in COVID-19 (20-23). The relative expression of these genes and their interaction according to each individual MEL-Index score were analyzed by three different bioinformatics approaches. First, we determined the correlation between MEL-Index and all the 455 genes for detecting the significantly positively and negatively expressed. Then we used the GSEA technique, also known as functional enrichment analysis, to discriminate processes related to variation in lung melatonin content. Third, we evaluated the degree of expression and connectivity segregated according to low, medium, and high MEL-Index scores. Considering these three different bioinformatics approaches, we constructed gene sets for evaluating the correlation of MEL-Index to genes linked to virus invasion and infection.
Analyzing all the 455 genes we observed that up to 47% of the genes were not directly correlated with the MEL-Index score and that the percentage of negatively correlated genes was higher (36%) than that of positively correlated ones (17%) (Figure 1, Appendices Table S1). The gene set enrichment analysis of the COVID-19-Signature confirmed that most of the genes negatively correlated with the MEL-Index, strongly reinforcing that normal lung melatonin production could be relevant in dealing with the virus invasion.

Considering that melatonin orchestrates different cellular and intercellular processes and its role changes according to cell phenotypes, we used a bioinformatics tool that integrates the connectivity between the most expressed genes, allowing us to compare the same set of genes under different situations. Thus, instead of evaluating the most or less expressed genes or enriched gene sets previously curated, this method discloses cohesive subgroups of variables in one of the states and evaluates whether these groups change their correlation patterns among states (26). The method is found useful to identify influential spreaders in a network, as instead of considering only the most highly connected or the most central genes, it considers those that are located in the core of the network, serving to different levels of organization (32). This theoretical approach firstly confirmed that MEL-Index scores define different network states in normal lung, and secondly shows that in the High MEL-Index group, the density of the 40 most expressed genes was reduced, and the connectivity among them inhibited. Thus, the pattern of connectivity among genes that integrates different networks is inverted when low and high-MEL-Index are compared (Figure 3A), suggesting that the presence of the virus is less noticed in lungs that synthesize melatonin. At this point, we decided to explore networks that could be the biological basis for virus invasion.

The gene networks chosen to be related to the MEL-index score involved virus gateway (receptors), intracellular traffic, transcription/post-translational control, and mitochondrial function. Genes involved in the disease, such as immune responses and lung dysfunctions, were not included in our panel. The data obtained strongly suggest that lung melatonin content could be a protective factor against virus invasion. Indeed, more than 80% of the top 30 MEL-Index negatively correlated genes were associated with virus gateway (3 genes), transcriptional/(post)translational control (11 genes), intracellular traffic (7 genes), or mitochondrial functions (4 genes). Among the positively correlated genes, two of them are transcription repressors (ZNF503 and TLE1), and one controls mitophagy (USP54), a critical process for mitochondrial quality control (33). As a matter of fact, melatonin is known to attenuate anoxia/reoxygenation injury by inhibiting excessive mitophagy (34). TLE1, positively correlated in virus gateway and transcription/translation sets, encodes a repressor transcription factor for the expression of pro-inflammatory NFκB subunits, being an inhibitor of the mounting of innate immune responses (35).

Mitochondria is a pivotal player in orchestrated defense responses, besides efficiently converting glucose into ATP. The role of mitochondria varies according to cell phenotype and changes in the cytoplasmic and extracellular microenvironment, including regulating the response to viruses (36). In COVID-19 mitochondrial ROS/hypoxia-inducible factor-1A (HIF-1A)-dependent pathway activation results in dysfunctional immune response and epithelial cell death (37). Otherwise, melatonin is known to reduce HIF-1A expression and stabilize mitochondrial function (38). Here we observed that HIF1A was one of the most negatively correlated genes with MEL-Index score (r -0.37; p 1.2E-10; FDR 5.2E-8). Besides HIF-1A, genes that codify proteins linked to catabolism (BCKDK), acidification of intracellular compartments (ATP6V1C1, ATP6V1A, ATP6V1D), electron chain transfer (NDUFAF2, ETF), transport across mitochondrial membranes (GRPEL1, TIMM8B, TIMM9, TIMM10), transcription (MRPS5), organelle biogenesis (STOML2), and cytoprotection (HYOU1, GPX1) were also negatively correlated with MEL-Index (Appendices Table S1). In several cases, overexpression of these genes may result in mitochondrial or cellular damage, and therefore, it
supports the hypothesis that the High-MEL-Index score could confer higher resistance to infection.

A similar general profile was observed for genes linked to transcription/post-translation modulation and cellular traffic. Genes linked to RNA processing and degradation (EXOSC3, EXOSC5), ribosomal assembly (HEATR3), endoplasmic reticulum folding, and assembling of proteins (HSPA5) and secretory pathways (UPF1) were also negatively correlated. Besides, genes codifying proteins linked to ribosomal integrity (LARP1, PABPC1) were also shown to regulate dengue virus replication (39-41).

After virus entry, the traffic inside the cell is adapted for optimizing virus replication and release. The 35 genes associated with intracellular traffic that negatively correlated with the MEL-Index (Figure 4), codify proteins that regulate trafficking in organelles and transport vesicles (RAB18, GRPEL1), protein transport from the endoplasmic reticulum to Golgi (RAB2A), vesicular transport (RAB 7A, VTA1, TMED5), lipid transport from the ER (SIGMAR1), transport of early endosomes (FKB15), nuclear transport (NUP62 and NUFT2). Besides this general approach, it was noteworthy that FURIN (r = -0.15; p<0.01) codified protein is involved in the cleavage of the SARS-CoV-2 spike (42). This gene gain relevance over the others; it selectively cleaves the SARS-CoV-2 spike, but not the SARS-Cov or MERS-Cov that use a similar gateway (43).

In summary, a High-MEL-Index lung score negatively regulates the expression of genes that facilitates virus mobility inside the cell. SARS-CoV-2 gateways in epithelial AT2 cells and alveolar macrophages occurs via different receptor complexes. The gene expression that codifies each cell's receptors (ACE2 for epithelial AT2 cells, and CD147 for alveolar macrophages) in the lung of healthy subjects present low or no variation with MEL-Index scores (Figure 5, Appendices Table S3). The complex of protein that forms the gateway in epithelial AT2 cells, but not in macrophages, was shown to be responsible for the infection (27). Thus, we will explore the relationship between MEL-Index and the other proteins that form the gateway complexes.

The ACE2 transmembrane protein binds to the virus spike in two stages. The intact spike binds to site one, and then after the cleavage of this protein by a transmembrane serine protease coded by TMPRSS2, the spike links to a second binding site attaching the virus to the plasma membrane (44). When the cleavage of the spike protein by the serine protease occurs before the linking to the first binding site, virus invasion is impaired (45). Moreover, another protease, now similar to trypsin, cleaves the SARS-Cov2 spike in a third site, mandatory for virus internalization (46). The presence of a furin-like cleavage site in SARS-CoV-2 facilitates the S protein priming and might increase SARS-CoV-2 infection's efficiency compared to other beta coronaviruses. So, furin inhibitors can be targeted as potential drug therapies for SARS-CoV-2 (47). The relevance of both proteases for SARS-CoV-2 infection is reinforced by a Pharmacological approach, as the antiviral activity of combined TMPRSS2 and furin inhibitors was more potent than the sum of equivalent isolated doses (43). As the MEL-Index score negatively correlates with the expression of FURIN and positively with that of TMPRSS2, the entry of the virus in AT2 epithelial cells will be hampered in High MEL-Index subjects, probably turning them into asymptomatic carriers. The data also reinforce the idea that treating the patients with melatonin via nasal installation or drops could reduce the rate of infection.

Nerve-associated, lung-resident interstitial macrophages (NAMS), identified in 2020, and alveolar macrophages are resident cells, yolk-sac derived, constituent the first line of defense (48). NAMS is localized in the nose and trachea, associated with sympathetic neurons. Another gateway mediates virus entry in resident macrophages, and internalization did not promote proliferation (examples: 1918 H1N1 influenza virus (49), respiratory syncytial virus (50), the PR8 influenza virus (51), and the Newcastle disease virus (52). Although no single-cell study could suggest a melatonin-related reduction in SARS-CoV-2 proliferation in alveolar
macrophages, our data clearly show that the proteins that compose the CD147 multi-molecular complex, the gateway for SARS-Cov-2 in macrophages are negatively correlated to MEL-Index scores (Figure 5, Appendices table S3).

**Fig. 5. Lung MEL-Index orchestrating SARS-CoV-2 invasion and infection.**

Left side - Healthy lung (A) showing an alveolus and the alveolar macrophage membrane (left) and epithelial AT2 cells SARS-CoV-2 receptors. Right side - of the figure the reaction of healthy low (B) and high (C)-MEL-Index people to SARS-CoV-2 invasion. In the epithelial AT2 cells, the virus binds to one site in (1), and then it is cleaved by (2), favoring the binding to the second site and internalization. High-MEL-Index scores increase the expression of (2). If the spike is cleaved before the first binding, the entry is hampered (see text). High-MEL Index also negatively correlates with the proteins that form the multi-molecular complex CD-147 receptor. (1)ACE2 - angiotensin-converting enzyme 2; (2) TMRSS2; transmembrane protease, serine 2 (3) CD147 - extracellular matrix metalloproteinase inducer or basigin; (4) MCT-monocarboxylate transporters & GLUT1 - glucose transporter 1; (5) CYP A – cyclophilin A; (6) CYP B cyclophilin B; (7) S100A9; (8) hyaluronic acid; (9) CD44 – also known as phagocytic protein 1. – The figure was adapted from (23), created with BioRender.

CD147 (basigin, BSG gene), a transmembrane protein member of the immunoglobulin superfamily, is the center of a multi-molecular receptor complex recognized as the virus gateway in renal, cardiac, gastro-intestinal, and alveolar macrophages (23). The direct interaction of the capsids of corona and HIV-1 to cyclophilin A mediates the binding with CD147 and cell entry (53). Cyclophilins A is a potent leukocyte chemoattractant, and its interaction with CD147 is studied in many contexts, including inflammation, immunological responses, and cancer (54). Contrasting with ACE-1, taken as the gateway of SARS CoV-2 in the lung, CD147 multi-protein complex receptors were implicated in COVID-19 pathogenesis, mainly in patients with comorbidities (23). A paper published in December 2020 reinforced this hypothesis (55). The expression of cyclophilin A in biopsied samples from patients with COVID-19 was enhanced in podocytes and parietal cells compared to non-COVID-19 autopsied tissue or acute kidney injury biopsies. It is noteworthy that, until now, there is no confirmation that concanavalin A immunosuppressed patients constitute a group of risk for COVID-19 (55). Concanavalin A induces immunosuppression via binding with cyclophilin A (56, 57) and blocking the downstream SARS-CoV-2 replication pathway, indeed the same pathway that leads to immunosuppression (58). In summary, there is a growing sum of
information pointing to the multi-molecular CD-147 receptor complex as an essential target for the entry of SARS-CoV-2 in other cells, including the alveolar macrophages (Figure 5).

Taking together the effect of AT2 and resident macrophages gateways, High Mel-Index will block infection, as the entry in AT2 epithelial cells is hampered, and will favor transmission, as resident macrophages gateways are also not expressed. Thus, the present work discloses a biomarker for virus transmission by asymptomatic carriers providing a biological tool for controlling the pandemic. Besides, strongly suggest that "healthy" positive subjects to SARS-CoV-2 mRNA should be preventively treated with nasal administrated melatonin, avoiding the evolution of presymptomatic carriers.

ACKNOWLEDGEMENT

The authors gratefully thank the financial support from the São Paulo Research Foundation (FAPESP) provided to PACMF (2015/23348-8), RPM (2013/13691-1), MSB (2008/57908-6 and 2014/50884-5) and SMM (2018/24693-9); the National Institute of Science and Technology of Bioethanol (INCT- Bioethanol) and the National Council for Scientific and Technological Development (CNPq: 574002/2008-1 and 465319/2014-9, provided to RPM). The founders had no role in the study design, data collection/analysis or preparation of the manuscript. Figure 5 was kindly designed by Débora PM Lotan and Sofia Lotan.

AUTHORSHIP

PAF and RPM concept and designed the work; PAF, GSK, RPM, BVN, VCJ, EDRP, MOCM, DSS, SMM collaborated in data acquisition; PAF, RPM, GSK, BVN, VCJ, AF, HIN and MSB collaborated in data analysis and interpretation; PAF and RPM wrote the manuscript, MSB revised and all the authors approved

CONFLICT OF INTEREST

Authors declare no conflict of interest.

REFERENCES


Please cite this paper as: