

SARS-CoV2 seasonality and adaptation are driven by solar activity.

Octavio Gonzalez-Lugo\*

<sup>\*</sup>Corresponding author *No affiliation*. Email: ogonzalezl@outlook.com

## Abstract

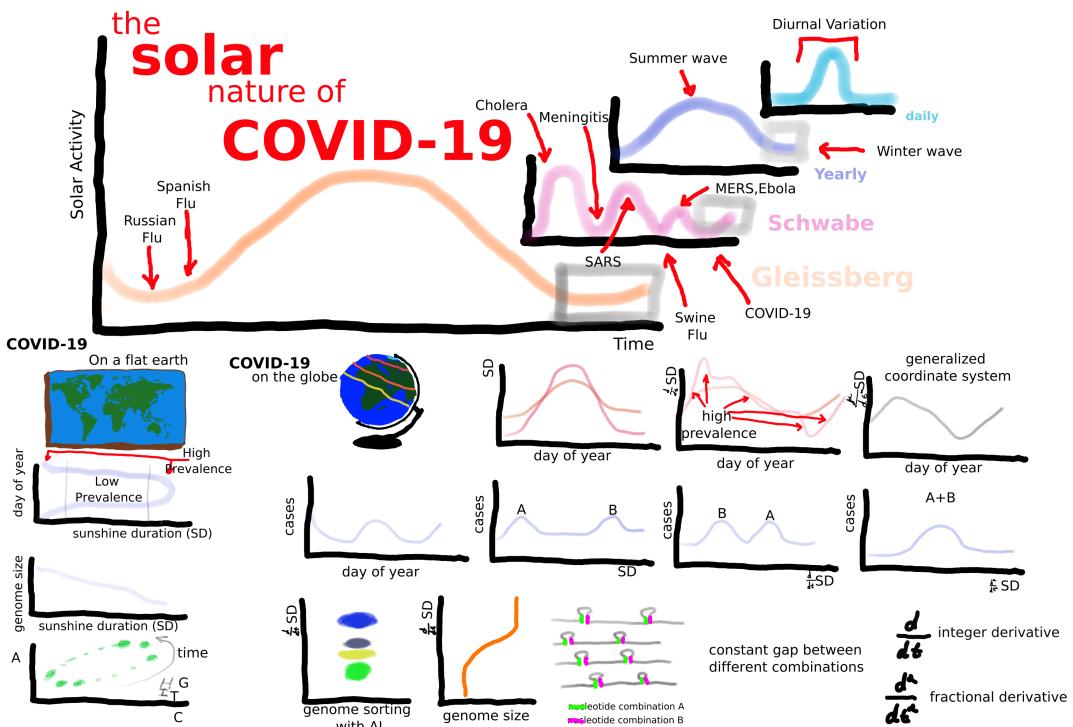
Since its isolation, SARS-CoV2 showed a high mutation rate hindering the ability to properly characterize it. Also as a consequence of its size, traditional sequence analysis methods are computationally constrained. Nonetheless, machine learning technologies and particularly deep learning offer a viable and scalable solution to the analysis. Applying variational autoencoders (VAEs), a special kind of neural network used for unsupervised learning, to custom genomic sequence representations cluster the genomes by similarity. The resulting learned representation sorts the SARS-CoV2 genomes by the sunshine duration (SD) rate of change (SDRC) and other solar-derived features. The transition between clusters is characterized by changes in viral genome size. Different deletions can be found throughout the SARS-CoV2 genome. Deletions might behave as an internal clock inside the genome. Or as an adaptation mechanism to seasonal variations in solar radiation. As genome adaptation follows SDRC COVID-19 cases could also follow the same pattern. However, the spherical geometry of our planet adds a confounding variable to the analysis. Spherical correction by dynamic fractional differentiation of the SD synchronizes COVID-19 cases into a single period. This analysis results in a general coordinate system that shows the seasonality of the disease and allows us to analyze the impact of other environmental features. Atmospheric changes that also affect solar radiation such as the ozone column and cloudiness gradients also showed a correlation with COVID-19 cases.

Analyzing viral genome composition (amount of each specific nucleotide) as time series displays an attractor-like behavior under different solar-derived time scales. Oscillations in genome composition could explain the changes in the efficacy of different antiviral treatments through the pandemic. To further characterize and expand the prediction capability of the model a set of solar and environmental features are added as targets in a supervised learning task. Allowing to correlate environmental changes and genomic composition.

Genome sorting is obtained in a two-dimensional learned representation and appears that only one dimension encodes a temporal variable. The learned representation could also show the interaction between the host or a particular tissue of the host and the virus. To further characterize the learned representation the VAE model is modified to also fit a predator-prey model using the learned representation as observations of the model. Retrieved dynamics shows a transitory phase with a duration similar to the SARS-CoV2 incubation time. However, further characterization is needed to accurately interpret those results.

Under a general coordinate system, the different mutations or genomic adaptations follow a regular and deterministic pattern. If the adaptations are the result or the means to copy the codon usage bias(CUB) of the host it is likely that SARS-Cov2 follows the expression of compositionally similar genes. Selection of highly compositionally similar transcripts results in the isolation of a series of genes whose dysregulation matches some of the symptoms reported in post-acute COVID-19 sequelae or long Covid. Also, genes that regulate the circadian rhythm and that are highly expressed at periods of high COVID-19 prevalence. Copying or optimizing genomic composition towards the host's CUB will increase the probability of the virus to copy a significant fragment of a host protein leading to autoimmunity. Dysregulation of the solar-based regulation of the circadian rhythm could explain the variability and random-like behavior of the symptoms in the post acute phase.

Further development of analysis techniques will help us to better understand the seasonality and adaptation of pathogenic organisms.



**Figure 1: The solar nature of COVID-19** The seasonal pattern followed by SARS-CoV2 is characterized by nonlinear changes through different time scales. The oscillatory nature of the environmental feature as well as the geometrical characteristics of our planet adds a series of confounding variables to the analysis. Under a general coordinate system, SARS-CoV2 follows a stable pattern of transmission as well as mutation/adaptation. The stochastic nature of mutations under a general coordinate system seems less than previously predicted, perhaps even absent.

49 **Introduction**

50 The ongoing COVID-19 pandemic generated a large quantity of data as a result of tracking both the  
51 virus and the disease cases. Those data resources were published regularly and openly throughout the  
52 pandemic. However, the size of the datasets restricted the ability to analyze the data to a few centers  
53 with high computational resources. Such a barrier is particularly true for genome sequence data.  
54 As the genome size hindered the ability to use large samples size to understand the evolution and  
55 adaptation of the virus [1]. The development of new sequence representation schemes as well as the  
56 application of deep learning methods to the new datasets could fast-track the characterization of new  
57 emerging or neglected pathogens. The following describes the use of different custom SARS-CoV2  
58 sequence representation schemes to train VAE models. And how to combine different information  
59 with the insights obtained by those models. Continuous development and improvement of sequence  
60 representation schemes will lead to a faster and more accurate characterization of emerging or  
61 neglected viral pathogens.

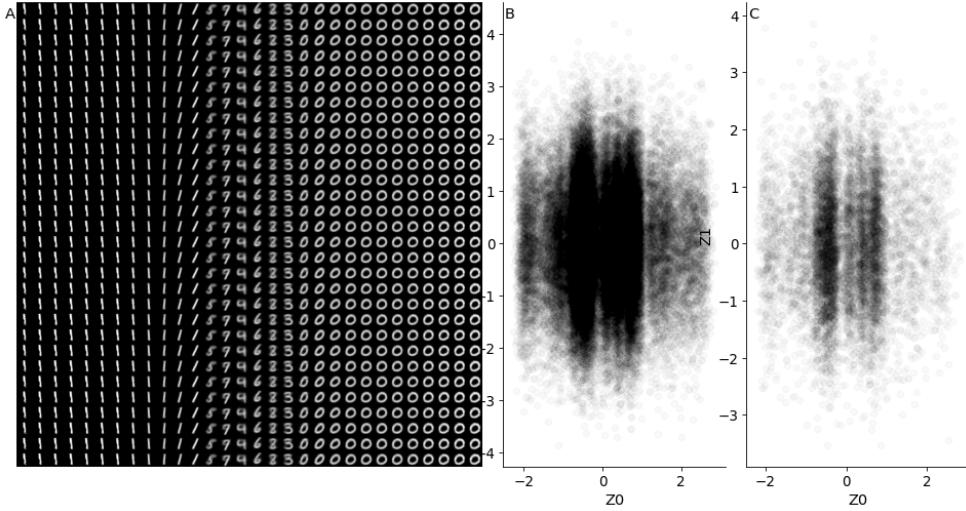
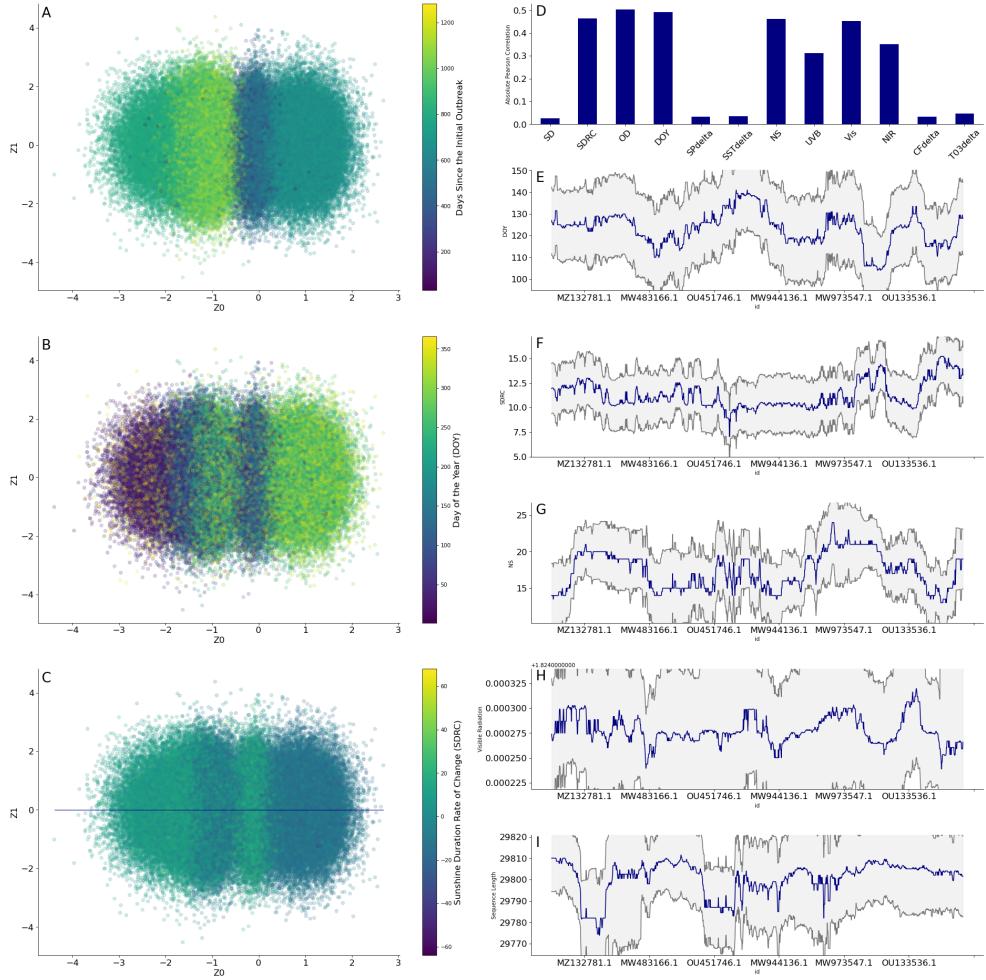


Figure 2: **VAE learned representation from the MNIST dataset.** A Visualization of the learned representation across  $Z_0$  and  $Z_1$ , the model learns the "roundness" of each number and aligns it across the  $Z_0$  axis. B, C learned representation, B shows the train data, and C the test data. Each dot represents a single sample from the MNIST dataset, encoding a 28x28 image into a 2D space.

62 **Generative models for sequence representation.**

63 Generative models are a special kind of machine-learning models that can generate new samples  
64 from a simple input. A VAE is one of the many different generative machine learning architectures.  
65 A VAE is a special network architecture, constituted by an encoder and decoder network, trained  
66 with the same data for input and output. As a consequence of the training, the encoder network  
67 finds a low-dimensional representation of the data that can be used for other applications [2]. When

68 there are no labels available for the data this kind of training is also known as unsupervised learning.  
 69 The low dimensional representation is also known as learned representation and each axis encodes  
 70 an attribute or pattern inside the data. Specifics of the attribute will depend on the data, the number  
 71 of samples, possible biases in those samples, and the model architecture. This allows for controlling  
 72 the behavior or attributes that the newly generated sample will have. (Figure 2)



**Figure 3: SARS-CoV2 encodes temporal information inside its genome.** A, B, C Latent space visualization, each dot represents a unique genome and colors the corresponding time scale. The line on C shows the range of genomes selected for further analysis. D Correlation between the Z0 latent dimension from the selected sequences and the different time-related features. E, F, G, and H Changes followed by the selected genomes under different time scales. I Changes in genome size through the found order. E, F, G, H, I 95% confidence interval in gray.

73 One of the main bottlenecks for sequence modeling is the size of the sequence, as the sequence  
 74 grows in size long-range dependencies start to be lost [3]. Thus designing different data structures  
 75 that better capture those dependencies or find other kinds of patterns inside the data is of great  
 76 interest. Furthermore, the application of generative modeling to genomic data, especially to fast-

77 adapting viruses, could fast-track the development of new treatments and the understanding of  
 78 emerging diseases.

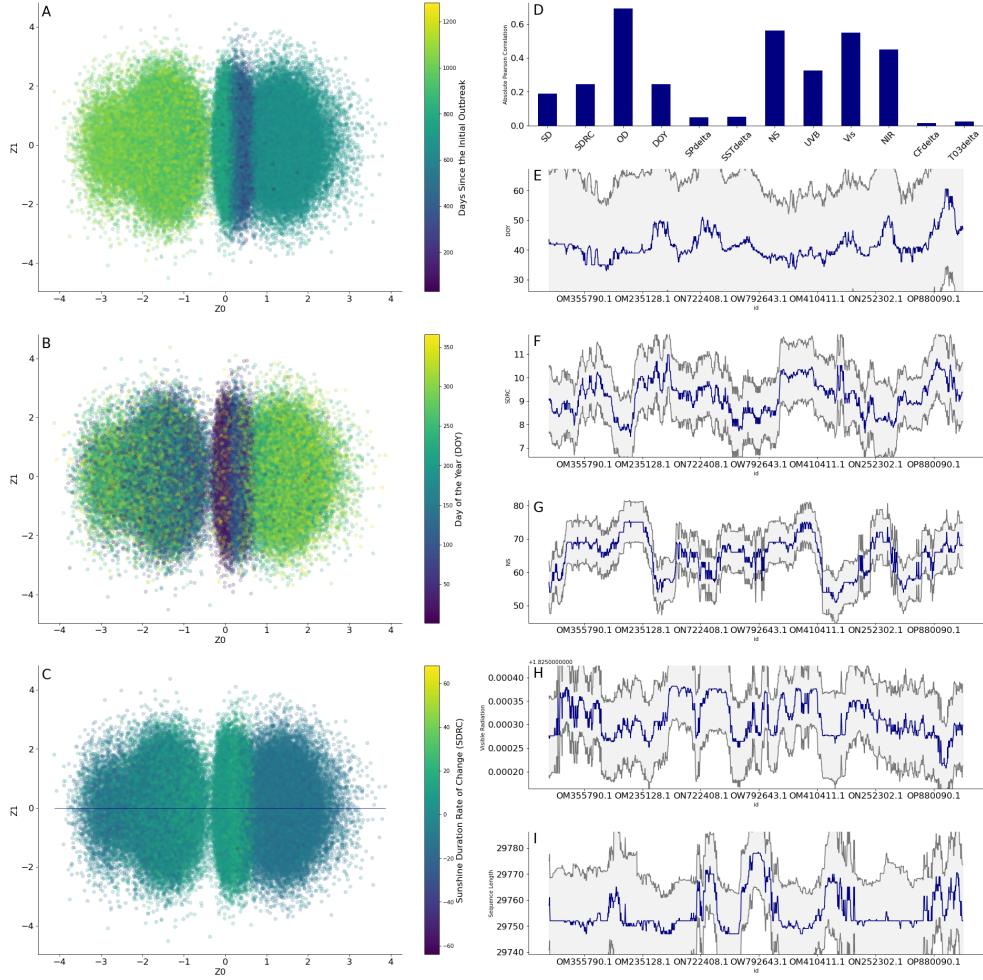


Figure 4: **Subsampling the genomic data biased the learned representation.** A, B, C Latent space visualization, each dot represents a unique genome and colors the corresponding time scale. The line on C shows the range of genomes selected for further analysis. D Correlation between the Z0 latent dimension from the selected sequences and the different time-related features. E, F, G, and H Changes followed by the selected genomes under different time scales. I Changes in genome size through the found order. E, F, G, H 95% confidence interval in gray.

## 79 Stacked K-mers

80 The first sequence representation consists of the normalized stacked frequencies of sliding sampled  
 81 K-mers. The sliding sampling scheme enables the analysis of different reading frames inside the  
 82 SARS-CoV2 genome. Trained VAE model results in a low dimensional representation that clusters  
 83 the sequences into a series of clusters ordered on one axis (Figure 3 A,B,C). Correlation analysis  
 84 shows that the latent dimension Z0 highly correlates with the SDRC the day of the year(DOY) as

85 well as different components in the solar radiation spectrum (Figure 3 D).

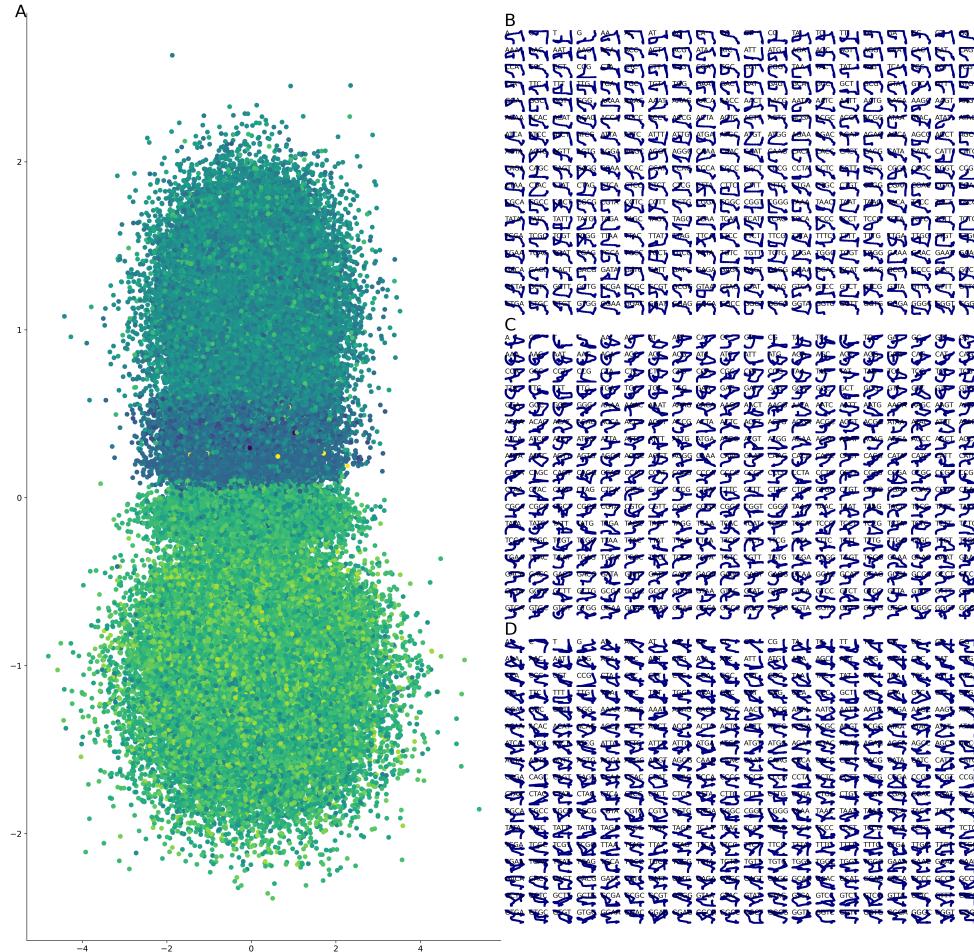


Figure 5: **Subsampling the genomic data biased the learned representation.** A, B, C Latent space visualization, each dot represents a unique genome and colors the corresponding time scale. The line on C shows the range of genomes selected for further analysis. D Correlation between the Z0 latent dimension from the selected sequences and the different time-related features. E, F, G, and H Changes followed by the selected genomes under different time scales. I Changes in genome size through the found order. E, F, G, H, I 95% confidence interval in gray.

86 To further characterize the information inside the latent dimension Z0 a series of sequences are  
87 selected close to zero in the latent dimension Z1. Changes in the SDRC and DY show an upward  
88 and downward trend similar to the pseudo color in Figure 3 E,F. While solar radiation features such  
89 as the number of sunspots (NS) and the irradiation of visible radiation follow a profile similar to the  
90 SARS-CoV2 genome size Figure 3 G,H,I.

91 However, genomic sampling is biased towards the second year of the pandemic and forward.  
92 This bias could influence the order found in the learned representation. To remove sampling bias  
93 as much as possible genomes are randomly selected to contain approximately the same number of  
94 genomes per day. Max number of genomes was adjusted around the peaks of the first and second

95 COVID-19 waves. The subsampled dataset also showed a high correlation towards solar-derived  
96 features. However, the high correlation towards SDRC is diminished several folds (Figure 4 D).

97 Changes in the correlation patterns point towards a minimum set of samples needed to shift  
98 the learned representation towards a specific space. The full dataset contains an excess of samples  
99 at the year level, thus the learned representation is biased towards patterns that better represent  
100 that characteristic. While the subsampled data contains fewer samples at a year level biasing the  
101 representation towards the long-term representation. Although specific changes related to a yearly  
102 or long-term adaptation have not been found yet. Changes in the learned representation hint at its  
103 existence.

104 With a better understanding of the environmental features that follow SARS-Cov2 to adapt  
105 its genome, the overall architecture of the network is modified to a supervised learning task. In  
106 this case, the model predicts a series of environmental features that highly correlate to SARS-Cov2  
107 adaptation as well as COVID-19 cases(Figure 12). Predictions obtained from the VAE are then  
108 used to predict the environmental features. This architecture resembles the function composition  
109 operation and allows the model to adapt the overall behavior to both representation learning and  
110 prediction.

111 Genome composition given by changes in UVB radiation pressure gradient and ozone gradient  
112 can be obtained by the model. However, no monotonous or simple pattern can be observed. (Figure  
113 5) Yet predictions of this model can be used to forecast the most likely composition of the genome  
114 given specific environmental conditions. This in turn could be useful to find an efficient and rational  
115 use of antivirals.

## 116 **Adjacency Matrices**

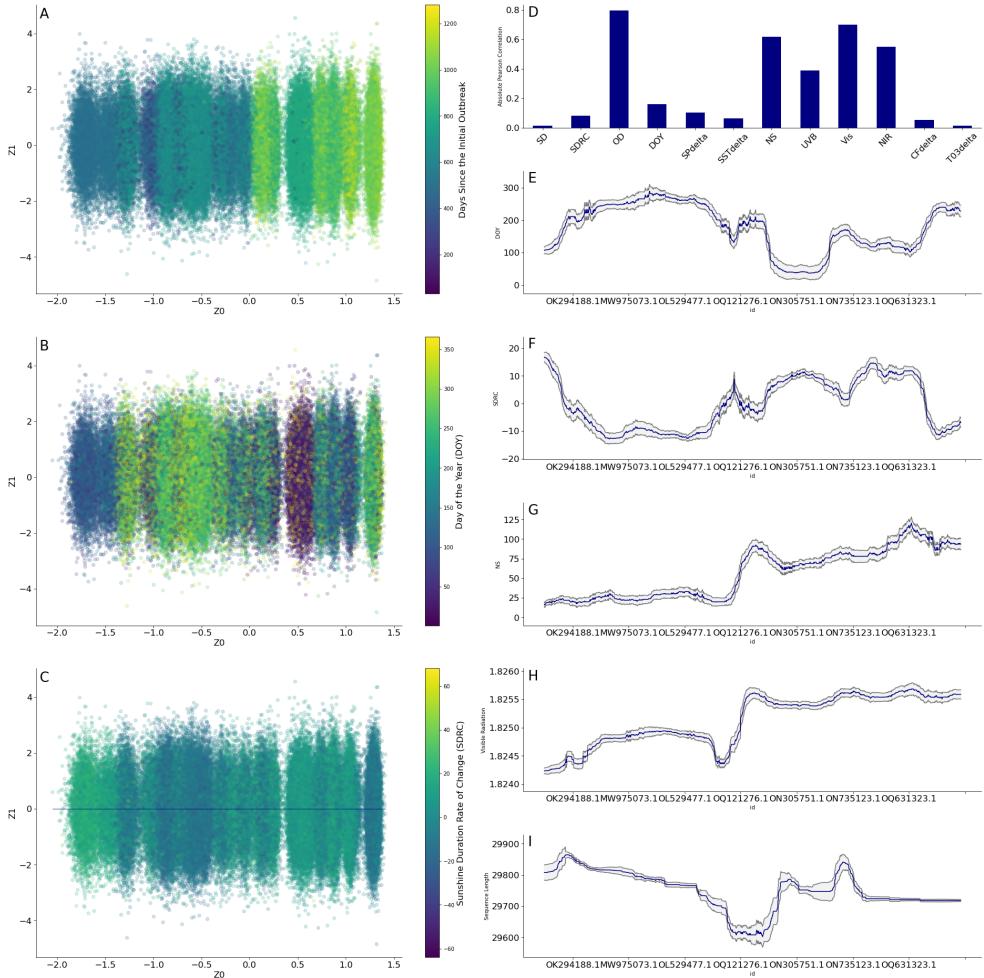
117 The previous dataset is capable to retrieve a specific correlation between the SASR-Cov2 genome  
118 and temporal scales, as well as other environmental features. However, it can't retrieve information  
119 regarding specific locations in the SARS-Cov2 genome. Therefore the design of different sequence  
120 representation schemes could ease the detection of such patterns.

121 The second dataset encodes the sequence by dividing it into 16 fragments and calculating the  
122 adjacency matrix of sliding consecutive 2-mers. This will encode the sequences by the connectivity  
123 between the different components rather than the frequency of each component. This will result in  
124 a hypercube of (16,16,16,1) dimensions.

125 Learned representation also retrieves a temporal pattern, although several features were tested  
126 no other significant correlation was found (Figure 6 D).

127 Yet the generative nature of the VAE model can also be useful to address how the genome changes  
128 through time and region. The decoder part of the model can be used to predict the changes along  
129 the learned representation. This kind of analysis is also known as a latent space walk. To ease the  
130 interpretation the input values for the random walk only differ along Z0 leaving Z1 equal to zero.

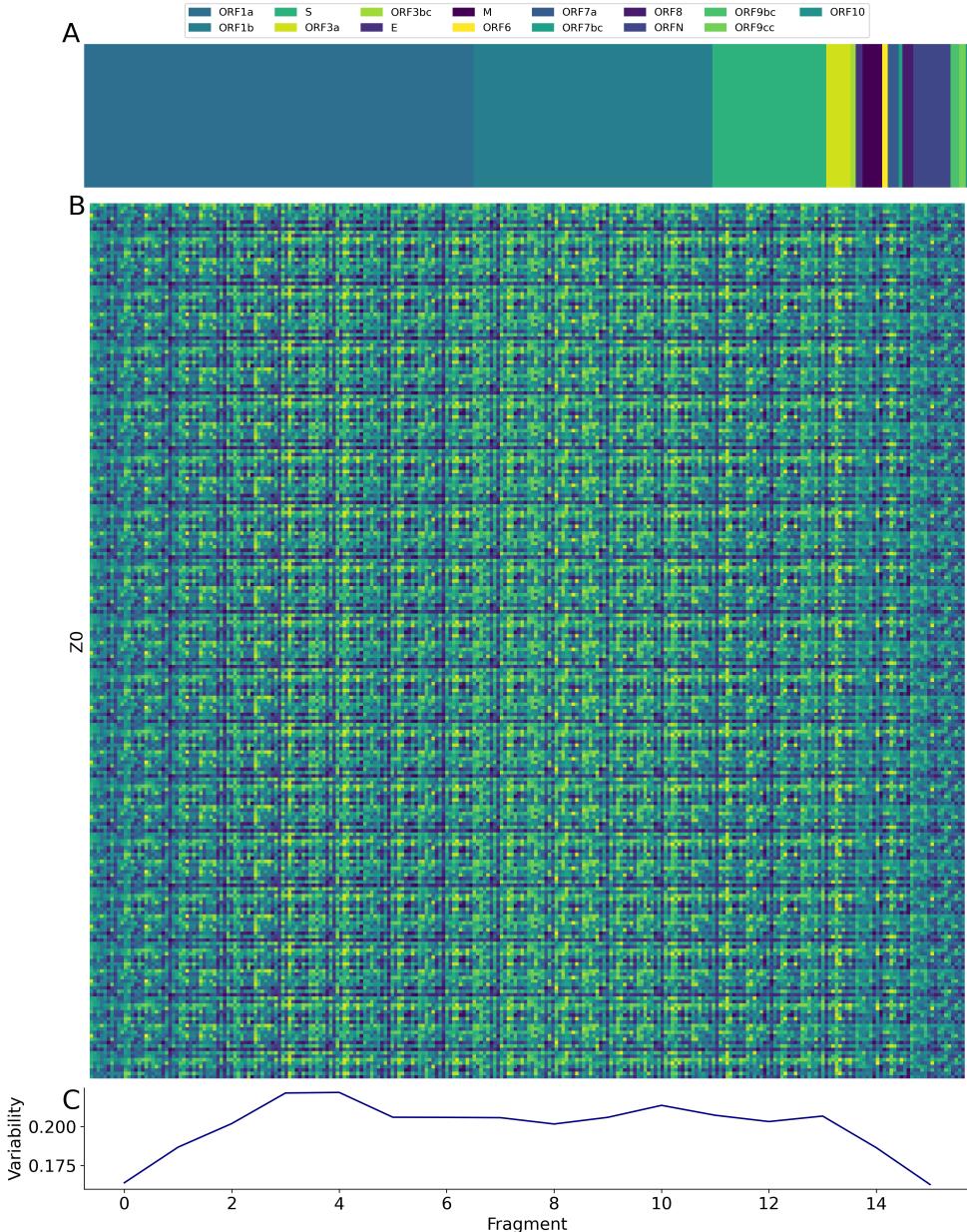
131 The standard deviation of the extreme regions of the SARS-Cov2 genome showed lower variability  
132 compared to the remaining regions (Figure 7 D). This characteristic of the genome could be used to  
133 target viral components in those regions as they will be less likely to change.



**Figure 6: Fragment-based representation better encode environmental features.** A, B, C Latent space visualization, each dot represents a unique genome and colors the corresponding time scale. The line on C shows the range of genomes selected for further analysis. D Correlation between the Z0 latent dimension from the selected sequences and the different time-related features. E, F, G, and H Changes followed by the selected genomes under different time scales. I Changes in genome size through the found order. E, F, G, H, I 95% confidence interval in gray.

Particularly at the 5' end of the genome Nsp1 and Nsp2 proteins could be potential targets due to the low variability in that region. Particularly Montelukast has been shown to inhibit Nsp1 by itself [4] or in combination with Ponatinib and Rilpivirine [5].

On the 3' of the viral genome both the nucleocapsid protein as well as the membrane protein align to the region with less variability. The nucleocapsid protein has been suggested to be added to new vaccine formulations as it is less prone to mutations [6]. Spike and nucleocapsid fusion protein using an adenoviral vector showed a specific response in animal models [7]. Using only the nucleocapsid protein for vaccination showed that inoculations of Lewis and Wistar rats with recombinant nucleoprotein resulted in no worrying side effects as well as the production of specific antibodies [8].



**Figure 7: Extremes of the SARS Cov2 viral genome are less variable.** A SARS Cov2 genome B Random walk around Z0, C Genome variability measured as the standard deviation of each fragment.

143 Regarding the membrane protein, inoculations with the membrane and envelope proteins in mice  
 144 showed partial protection [9]. Both combination or single nucleocapsid proteins could offer a better  
 145 long-term inoculation strategy due to low mutation in the nucleocapsid protein.

146 To further characterize the prediction capabilities of the model a new composition model is  
 147 trained using the previous genomic representation. As this new model can retrieve information

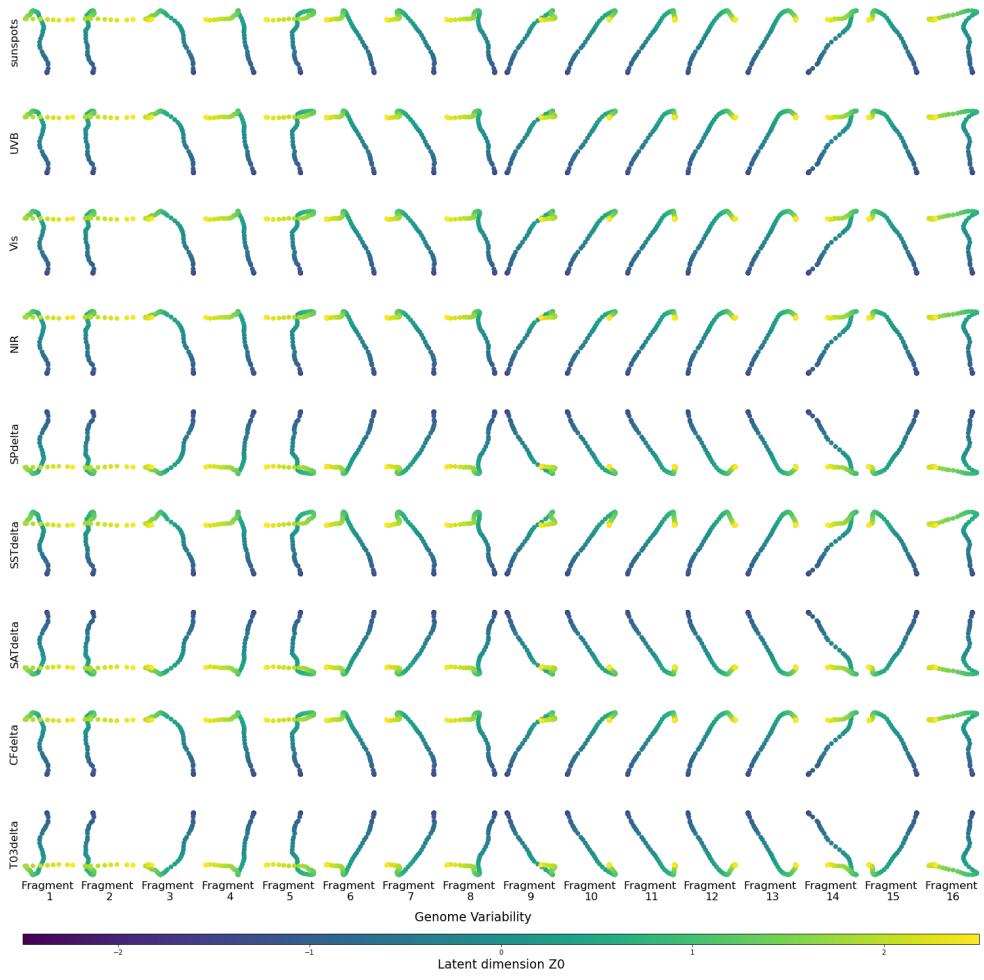


Figure 8: **Environment and genomic variability**

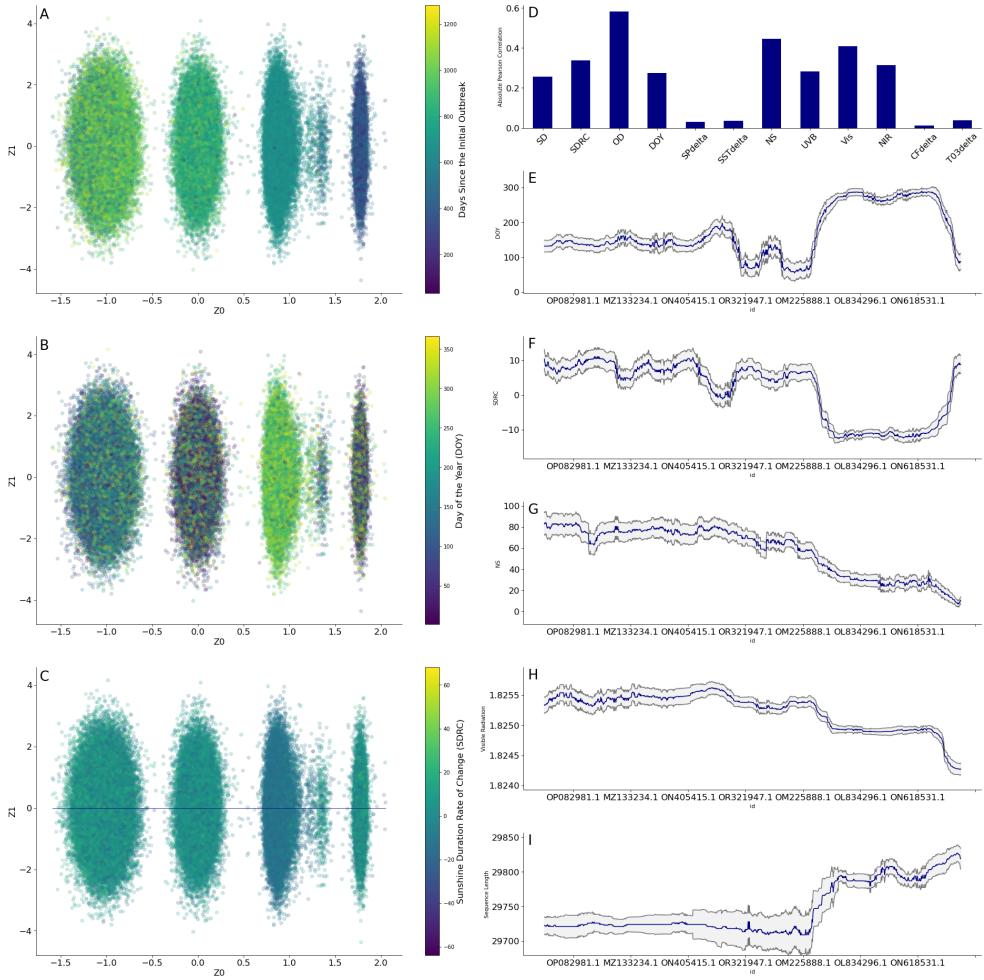
148 semi-location-wise, the trained model is used to evaluate the variability of the genome and its  
 149 dependence on environmental conditions. Although the accuracy of the environmental predictions  
 150 is low, a set of specific patterns can be retrieved. Particularly shifts in variability patterns along the  
 151 genome.

152 Changes in variability could point towards the genome converging into a specific configuration.  
 153 Convergence of variability patterns could forecast the kind of intervention that could be more suc-  
 154 cessful. Low variability in the non-structural coding region could suggest a higher success rate with  
 155 antiviral treatment. While low variability in the structural region will suggest a higher success rate  
 156 with vaccination or convalescent sera.

### 157 **Dimensional expansion.**

158 Previous models further confirmed the presence of a temporal pattern inside the SARS-CoV2 genome  
 159 and this pattern is highly correlated to solar activity. Either by the calculation of solar-derived

160 features such as the SD or SDRC as well as direct measurements of solar radiation at different wave-  
 161 lengths, or solar activity by the number of sunspots. They also showed the effect of data sampling  
 162 in the learned representation, as biasing the sampling towards a specific temporal characteristic will  
 163 result in a model that better represents such patterns. And that expanding the dimensionality of  
 164 the sequence representation could be a better strategy for large sequences. As the rearrangement  
 165 could bring closer different regulatory or similar regions mimicking the 3D structure of the genome.  
 166 To further test that hypothesis a third dataset was designed to predict the full-length SARS-CoV2  
 167 sequence. The dataset consists of the full length of one-hot-encoded sequences, and the resulting  
 168 encoding is reshaped to a hypercube of dimensions (32,32,32,4). Also, the number of samples is  
 169 lowered to reduce the yearly pattern influence and skew the results toward the long-range pattern.



**Figure 9: SARS Cov2 follows patterns of solar activity** A, B, C Latent space visualization, each dot represents a unique genome and colors the corresponding time scale. The line on C shows the range of genomes selected for further analysis. D Correlation between the Z0 latent dimension from the selected sequences and the different time-related features. E, F, G, and H Changes followed by the selected genomes under different time scales. I Changes in genome size through the found order. E, F, G, H, I 95% confidence interval in gray.

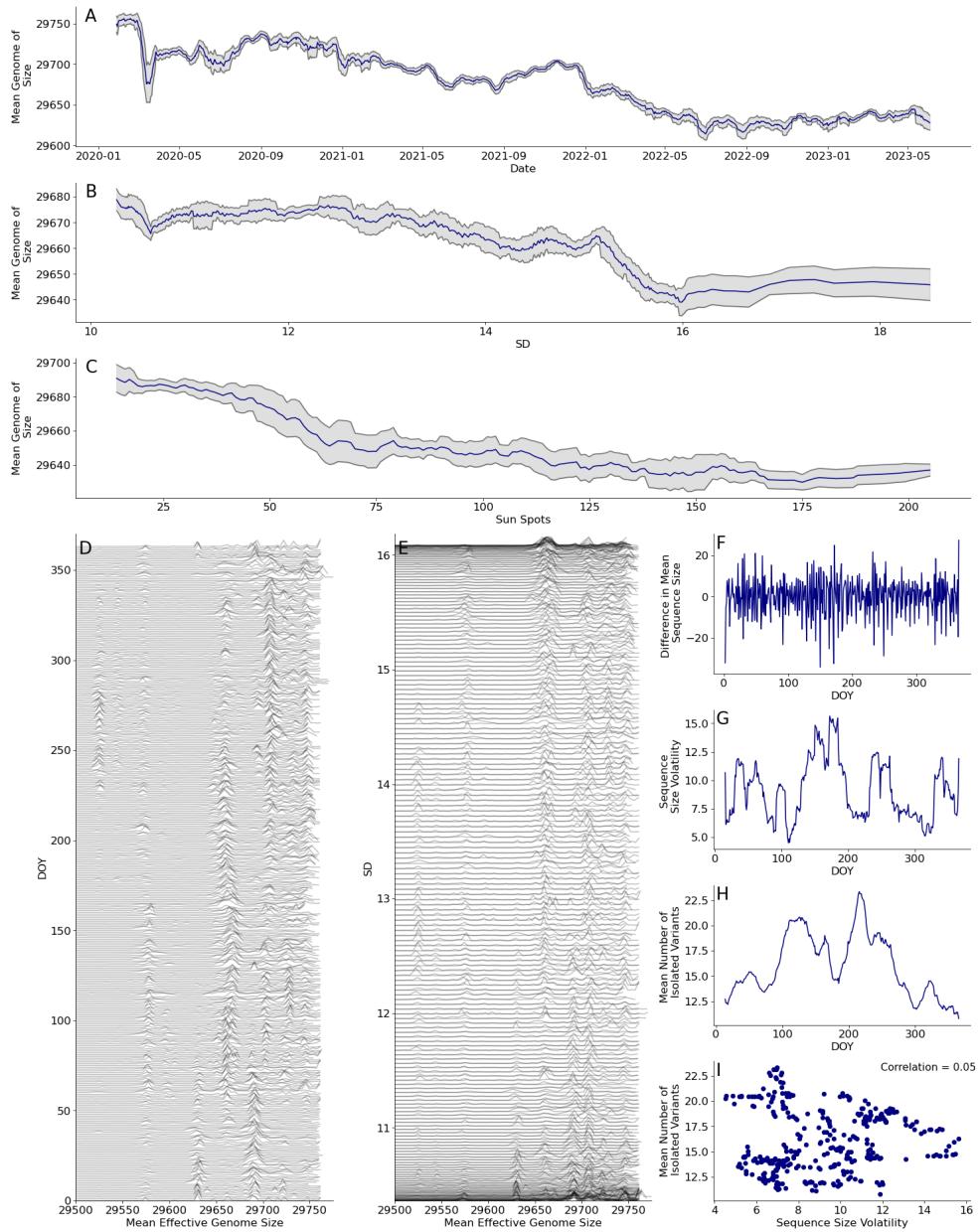
170 As expected the model can retrieve a representation that better fits the long-range pattern,  
171 however, is also capable to retrieve the yearly pattern encoded inside the SD and SDRC (Figure  
172 9 D). This multiplicity suggests a mixture of different time scales playing a role in SASR-Cov2  
173 adaptation and seasonality. A long-range dynamic, a yearly dynamic, and perhaps many others.

174 To test that hypothesis the genome size is used as a measurement of viral adaptation as genome  
175 size also follows the time scale inside the virus. Time since the initial outbreak shows a continuous  
176 decline in average genome size. The pattern is again repeated when the genomes are sorted by the  
177 SD or by the number of sunspots (Figure 10 A,B,C). This points to specific frequency components of  
178 solar activity, the yearly component encoded by SD, and the solar or Schwabe cycle [10]. A 12-year  
179 cycle that is measured by the number of sunspots. A third component will be the daily changes in  
180 solar radiation that vary throughout the day. SD measures the duration of such a period but not  
181 how solar radiation changes through the day. Genome size histograms at one specific location show  
182 a series of peaks that skew toward smaller genome sizes at periods with high SD and vice-versa.  
183 Furthermore, each day contains a series of distributions that might point to isolations at different  
184 times of the day (Figure 10 D,E).

185 Continuous rearrangement of the genome could be one of the many drivers behind the generation  
186 of new variants. Rapid changes in solar activity could drive the rearrangement of the viral genome  
187 leading to the generation of new variants. Measuring the genome rearrangement as the volatility of  
188 the mean genome size deviation and the genomic diversity as the number of variants shows a small  
189 negative correlation between both metrics (Figure 10 F,G,H,I). Although the correlation is small  
190 it could point toward specific mutational hotspots where the generation of variants is more likely.  
191 Those mutational hotspots could be repeated through the different frequency components of solar  
192 activity.

193 The combination of different frequency components could lead also to recombinant-like genomes.  
194 Identification of recombinant variants might point towards a genomic rearrangement derived from  
195 a low-frequency component within the solar cycle. The first isolation of omicron was on November  
196 2021 when the Schwabe cycle was entering a linear range. As the solar cycle was entering a new phase  
197 the transition could lead to a recombinant-like virus. First isolations of delta-cron a recombinant  
198 variant between omicron and delta were reported in January 2022 [11]. The linear phase of the  
199 Schwabe cycle is forecasted to end by 2024 NOAA [12]. Thus it's likely that more recombinant-like  
200 variants either with similarity to previously extinct clades or the generation of a novel clade will be  
201 isolated in the following months of 2023.

202 Genome rearrangements on the daily component have also been reported, although there were  
203 scarce at first. A Case report where the main infecting variant changed between samples taken  
204 at different times was published by [13]. Also, co-infections and genome variability started to be  
205 reported [14] [15]. Variability within the host might point towards a high-frequency component or  
206 the combination of different frequency components within the solar cycle rather than a co-infection  
207 scenario. Easement in the detection of the different genomic rearrangements could be the result of  
208 a higher adaptation to the codon usage bias(CUB) of the host. As the virus better adapts to the  
209 host's CUB fewer viral particles will be needed to infect the host, effectively increasing the R<sub>0</sub> of the  
210 virus. CUB adaptations could remain silent as a synonym codon will be favored over another [16].



**Figure 10: Effective genome size follows different scales of solar activity.** A Mean SARS-CoV2 effective genome size from Dec-2012 to Jun-2023, B Mean SARS-CoV2 genome size grouped by SD. C Mean SARS-CoV2 genome size grouped by the number of sunspots. D Changes in genome size by day at a single location. E Changes in genome size by SD at a single location. F Changes in mean genome size by day of the year. G Volatility or standard deviation of the mean change in genome size by day of the year. H Mean number of isolated variants at the location by day of the year. I Correlation between Genome size volatility and the number of isolated variants.

Dependence on solar activity could explain the benefit of different pharmacological and non-pharmacological interventions that mimic the action of solar radiation. Supplementation with vi-

213 tamin D has been found to confer protection against COVID-19 [17] [18]. Melatonin treatment  
214 also increases the clinical recovery rate [19]. IR radiation increases the recovery time of COVID-19  
215 patients [20]. These different interventions by themselves are not able to prevent SARS-Cov2 infec-  
216 tions. However, in conjunction, they show that correct circadian signaling is key for fast COVID-19  
217 recovery.

218 Although the different models are also capable to classify the viral genomes into several clusters,  
219 they do not match any given variant. The different clusters might represent viral pseudo species or a  
220 specific branch in the SARS-Cov2 phylogenetic tree. Different clusters also point to shared patterns  
221 among the different genomes in the same cluster. Although the specific nature of the pattern is not  
222 investigated its existence and classification could facilitate the design of new treatments. Recurrent  
223 mutation patterns could constrain the number of possible immune evasive variants or point towards  
224 a conserved mutation path. The lack of recognition of those patterns could be the result of the  
225 nonlinear nature of the viral adaptation or heavily relying on the Wuhan isolate or ancestral strain.  
226 The fast mutation rate of SASR-Cov2 also implies that the ancestral strain could be another varaint  
227 and not the initial virus that originated the outbreak.

228 Despite the fact that there are many examples of the application of machine learning models for  
229 SARS-Cov2 genomic data with even more sophisticated architectures [21] [22], the solar correlation  
230 has not been presented before. Perhaps the main reason behind that is how the model was used,  
231 rather than a tool for prediction, the model was used as a high dimensional sorting algorithm. Then  
232 different features were used to find the meaning behind the latent dimensions.

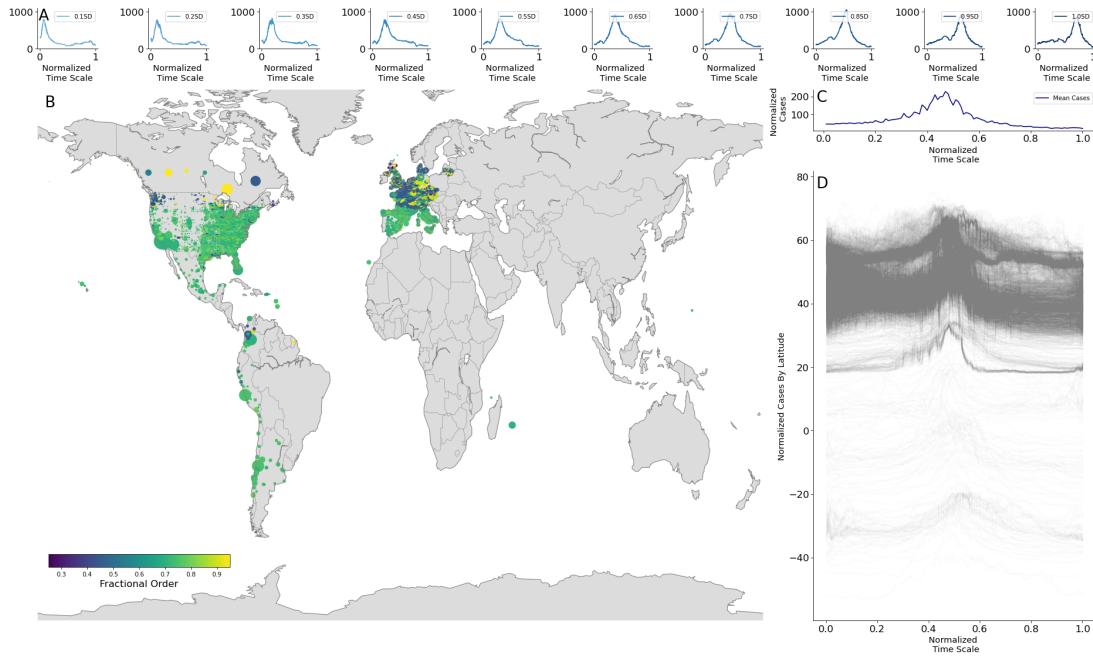
233 Sorting the data by autoencoders has been previously used to obtain meaningful coordinate  
234 systems from complex data sources [23]. This allows us to better understand the dynamics of a  
235 system by applying known dynamical models. Data obtained by genomic surveillance might not  
236 represent entirely the continuous adaptation of the virus as a series of recurrent ancestors. But  
237 rather a series of snapshots at different times of the interaction between a tissue within the host and  
238 SARS-Cov2.

## 239 **Epidemic curves under a generalized coordinate system**

240 Although there was a difference between folds and models the SDRC showed a high correlation,  
241 suggesting that it could encode the seasonality inside the COVID-19 waves. Displaying the number  
242 of cases with SD as a time scale for COVID-19 cases shows cases clustered at both ends of the time  
243 scale. While SDRC swaps the order of the peaks, suggesting that there could be an intermediate  
244 step between SD and SDRC that synchronizes both peaks into a single period. As SDRC is the  
245 first derivative of SD a smooth transition between SD and SDRC could be obtained by fractionally  
246 differentiating SD. Fractional differentiation is the generalization of the derivation and integration  
247 operation over fractional orders [24].

248 Fractionally differentiated SD (fSD) slowly synchronized the different peaks into a single one  
249 (Figure 11 A). Synchronizing the different COVID-19 waves by adapting the fractional order shows  
250 a relatively similar value except near the Equator and latitudes higher than 40 deg (Figure 8 B).

251 The fSD provides a general time scale to analyze the cases data regardless of the location.



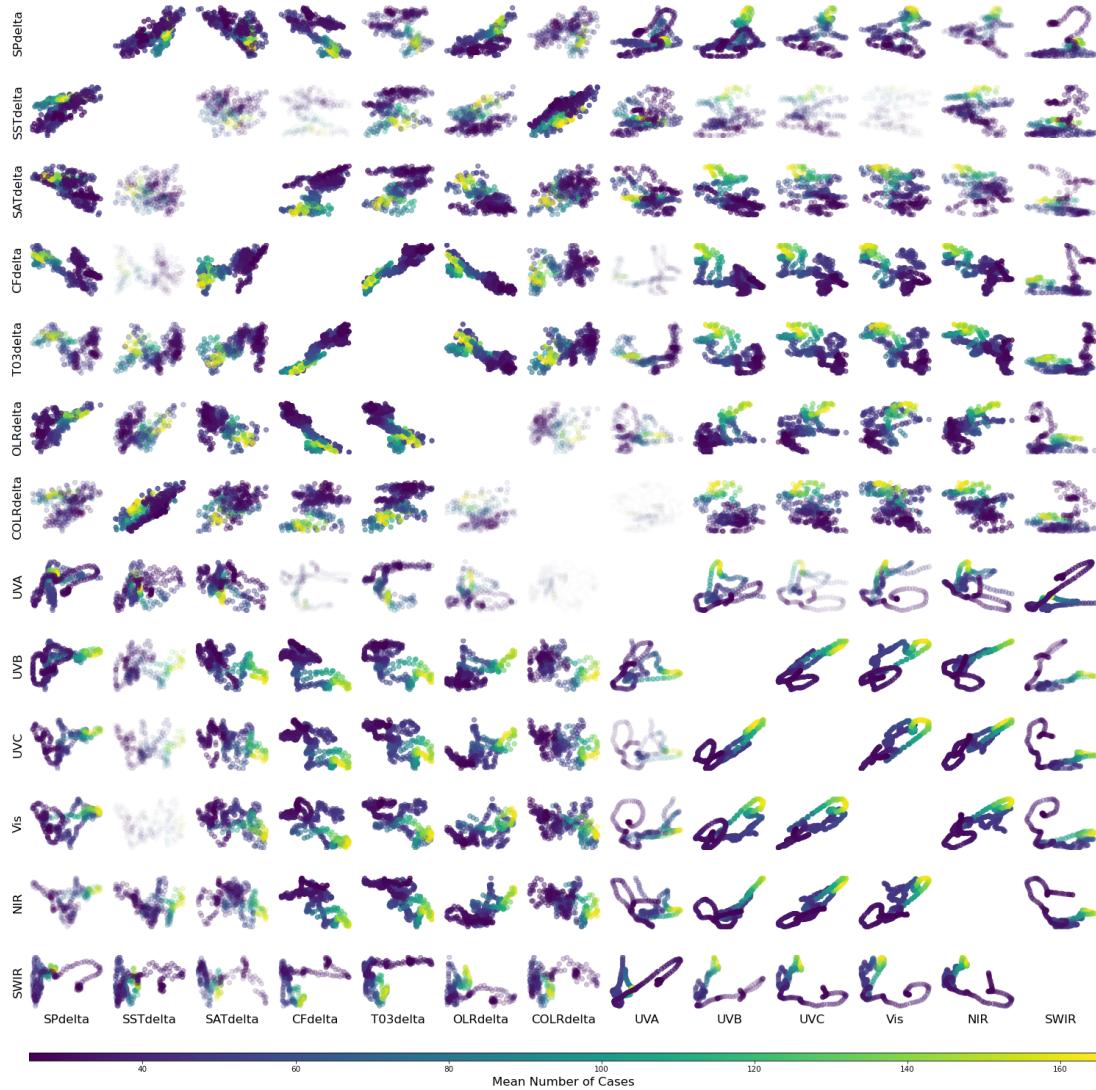
**Figure 11: COVID-19 pandemic dynamics follows the fractional SDRC.** A. Synchronization of epidemic curves by fractionally differentiating SD as a time scale. From left to right increasing order of fractional differentiation. B. Fractional order that synchronizes the epidemic curve at a specific geographic location. The marker shape is equal to the mean number of reported cases and the color is equal to the fractional order that synchronizes the epidemic curve. C. Mean synchronized curve. D. Waterfall plot of the synchronized curves by latitude.

The synchronization also allows to evaluate the correlation between cases and other environmental variables. Analyzing the mean number of cases over the generalized time scale shows high correlations between different wavelengths of solar radiation. Particularly a linear correlation is found between UVA and UVB and the number of cases of COVID-19 (Figure 9).

Environmental variables obtained from the Aqua/AIRS L3 Daily Standard Physical Retrieval dataset shows also linear correlations, particularly over environmental variables that influence solar radiation. Changes in the total column of ozone as well as changes in the cloud fraction. Ozone is an atmospheric component that absorbs the majority of UV radiation at the different wavelengths. While clouds, whose main component is water, absorb infrared radiation and UV radiation [25]. Correlation with other environmental variables are also found, yet are less in magnitude compared to radiation wavelengths and atmospheric variables (Figure 12).

Coordinate system discovery from the different models allowed the synchronization of the different epidemic curves at a city level. This specific data treatment will remove most of the socio-economic and behavioral effects of combining the data by country, age, or other proposed segmentation. It also removed the effect of outliers either due to high or low testing frequency, yet advantages of other kinds of segmentation were lost.

The relationship between solar activity and susceptibility to pathogenic diseases is a hypothesis previously proposed,[10] showed the correlation between different outbreaks through time and sta-



**Figure 12: Correlation between COVID-19 cases and environmental features under a generalized coordinate system.** Color represents the mean number of cases while brightness correlation between each feature. environmental features are calculated as the difference between the ascending and descending scans from the reconstructed dataset. SP (Surface pressure from forecast. (hPa)), SST (Surface skin temperature. (Kelvin)), SAT (Temperature of the atmosphere at the Earth's surface. (Kelvin)), CF (Combined layer cloud fraction. (0- 1). (Unitless)), T03 (Total integrated column ozone burden. (Dobson units)), OLR ( Outgoing long-wave radiation flux. (watts/m<sup>2</sup>)), COLR (Clear-sky outgoing long-wave radiation flux. (watts/m<sup>2</sup>)).

tionary points in the Schwabe cycle. Dependence between the susceptibility to SARS-Cov2 and the yearly change in solar radiation can be interpreted as a higher frequency component of solar activity. An even higher frequency component will be the daily changes in solar radiation that encompass the circadian rhythm. Dependence of COVID-19 on daily changes in solar radiation has been presented previously. Detection efficiency peaks at around the middle of the day [26] [27] [28]. Also, hospital

275 admission has been found to have a particular diurnal dependence as reported by [29].

276 As stated before correct circadian signaling is of extreme importance for COVID-19 recovery and  
277 protection. Stationary points of solar radiation such as those found in the middle of the year and  
278 winter, or at sunrise, sunset, or the middle of the day might impact the ability to signal those events  
279 if parts of the pathway are hindered either by environmental factors or host factors.

280 Another environmental factor that could impair circadian signaling, especially during the start  
281 of the outbreak, was the low solar activity. Initial cases of COVID-19 were reported at the end of  
282 2019 where sunspots were almost absent as we were entering solar cycle 25. If another frequency  
283 component of solar activity is affecting circadian signaling most likely will be of lower frequency.  
284 The Gleissberg cycle is a cycle with a duration of around 100 years [30], low levels of solar radiation  
285 at the Gleissberg frequency also correlate to the start of the COVID-19 pandemic as well as to other  
286 major outbreaks, particularly the Spanish Flu outbreak

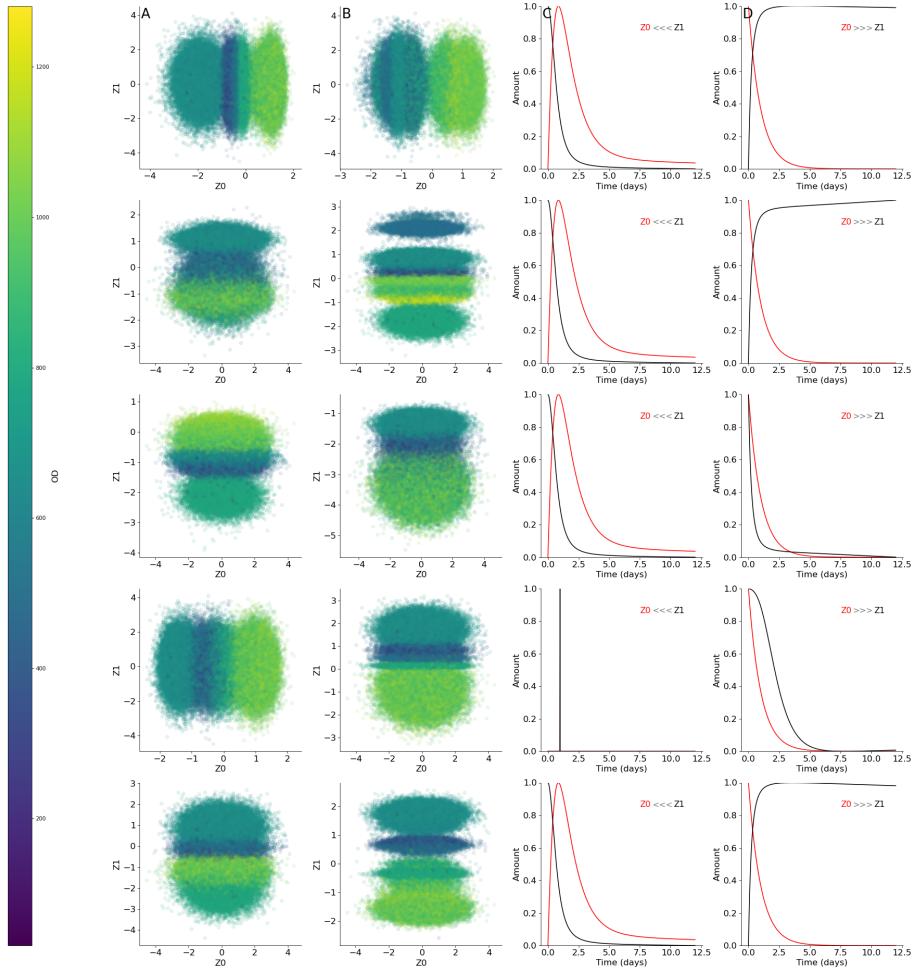
287 Conditions found at stationary points at the different frequency components of solar activity  
288 could impair circadian signaling, especially if the host is compromised over those pathways.

## 289 **Reframing viral genome reconstruction.**

290 Previous models suggest a deterministic mechanism behind viral replication and the role of random  
291 mutations most likely is less than previously suggested. How the virus replicates its genome and  
292 reconstructs the viral particle is an area of continuous research and very challenging. As the virus  
293 adapts to the host new strategies might evolve as a result. An example of such adaptation is the  
294 discovery of new entry pathways of SARS-Cov2 into the cell. The first entry mechanism was by  
295 the interaction between ACE2 receptors, while the second involved TMPRSS2 [31]. Viral entry  
296 and other phenotypes could be the result of the adaptation inside the viral genome and might be  
297 preferred for a specific cluster of SARS-Cov2 variants. How those different strategies emerge could  
298 be addressed by reframing the viral reconstruction process as different problems. They might share  
299 some properties among them but the methodology used to analyze them will differ between them.

## 300 **As a dynamical system.**

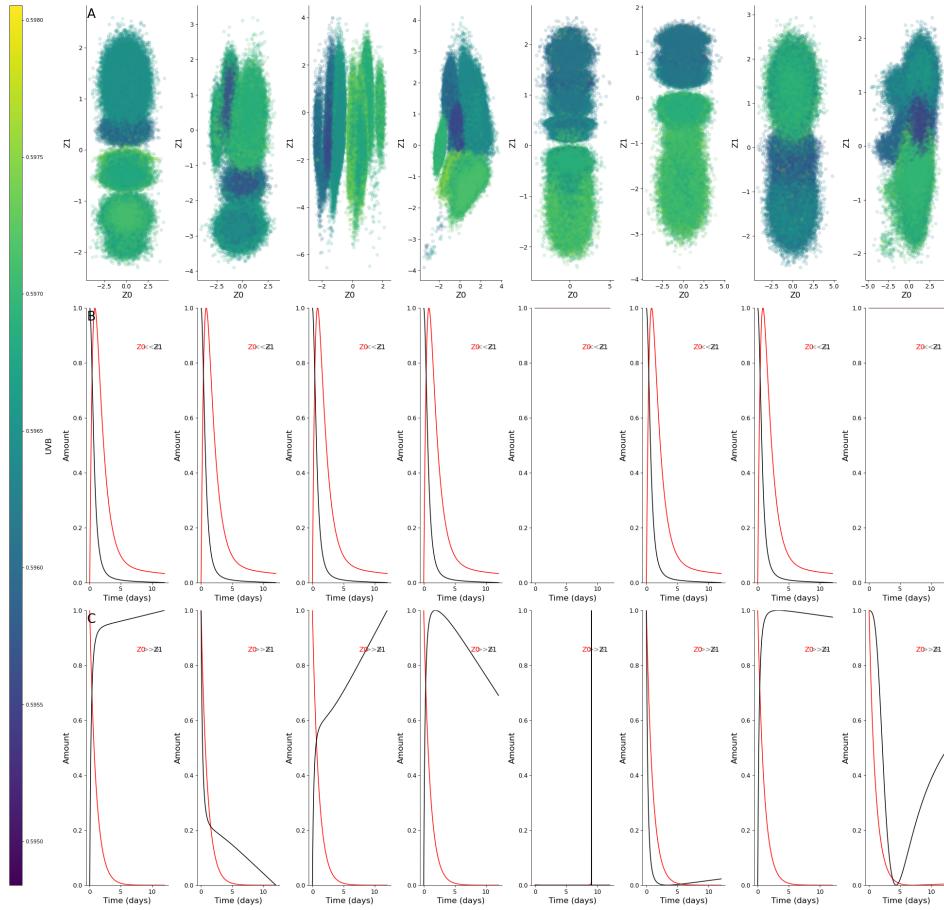
301 The learned representation obtained from the different models showed a cyclical pattern in one  
302 dimension, while the meaning of the second dimension remains to be addressed. One possibility  
303 of interpretation behind the different clusters and the learned dimensions is that they encode the  
304 interaction between the host and the virus. Each dimension might encode the resource availability  
305 in the host and the resource expenditure from the virus. This interaction could be modeled as a  
306 predator-prey model, where the available resources inside the host are the prey and the predator is  
307 the virus. One simple way to address such behavior is by modifying the VAE architecture to also  
308 retrieve a dynamical system. The learned representation is derived and the different constants of  
309 the predator-pray model are added to the optimization algorithm to be found. The use of autoen-  
310 coders for the detection of parsimonious dynamical systems has already been proposed as a feasible  
311 technique for complex data sources [23].



**Figure 13: Dynamical VAE stabilizes training and allows the clustering of SARS Cov2 genomes.** A Learned representation was obtained without dynamical loss, B Learned representation was obtained with a dynamical loss, same samples were used to train each model row-wise. D, E Retrieved dynamical system, initial condition displayed in the plot

312 Retrieved dynamics (assuming the time scale is in days) show a peak around the third day and  
 313 a duration of the transitory phase of around 5 days. This duration roughly matches the incubation  
 314 period of SARS-Cov2 [32] pointing to the importance of applying such models, specifically at the  
 315 beginning of an outbreak where information is scarce. Furthermore, adding the dynamic part to the  
 316 VAE stabilizes training. As VAEs trained using similar training conditions and the same samples  
 317 sort the viral genomes with less separation between clusters in comparison with their dynamical  
 318 counterparts (Figure 13).

319 Being able to approximate the incubation time from genomic surveillance data shows the impor-  
 320 tance of genomic data as well as one of the many applications of such data other than mutational  
 321 patterns. However, at the beginning of an outbreak, the amount of information is scarce making  
 322 it difficult to obtain such an approximation. Dividing the complete dataset into 8 different periods



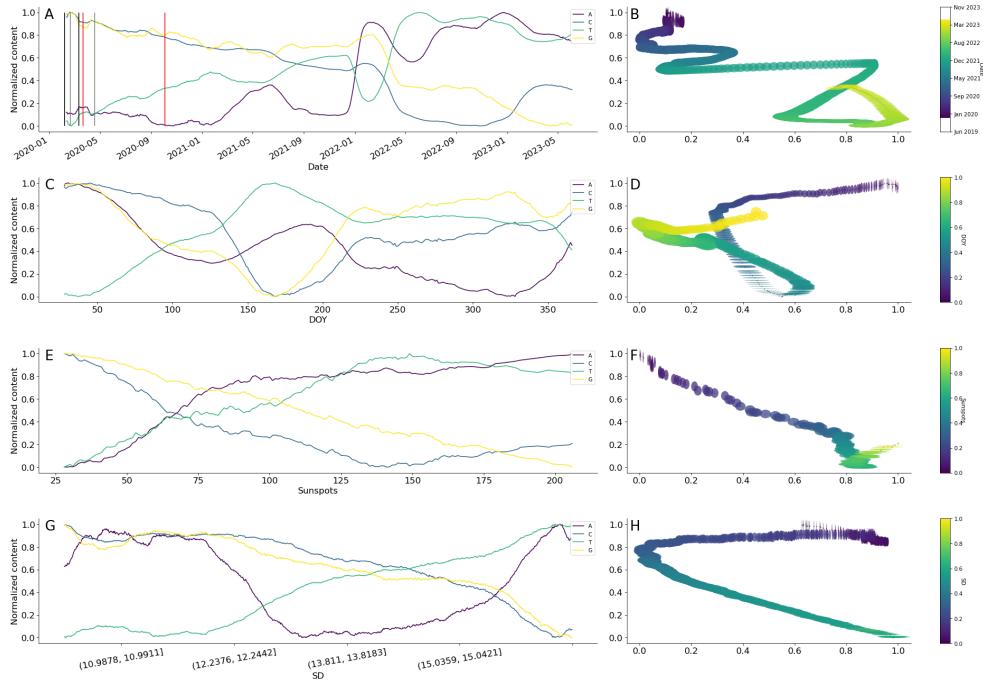
**Figure 14: Dynamics can be retrieved by a single fragment of the yearly cycle.** A learned representation was obtained by segmentation of the data into 6 consecutive periods by SD. B, C Retrieved dynamical system under different initial conditions.

using SD as a time scale shows that each model can retrieve a dynamical model with the same duration of the transitory phase.

Each period has a duration of about 1.5 months and if the samples are equally distributed around large geographical extensions the number of samples needed per location will decrease dramatically. Sharing data at the early stages of an outbreak of a novel pathogen could be the only way for such approximations to be obtained (Figure 14).

Nevertheless, caution is needed with the interpretation of these results. As most likely only represent the interaction with a specific tissue within the host, particularly the upper respiratory tract as is the tissue origin for most of the samples. Recurrent COVID-19 symptoms with a duration greater than the incubation time could be suggestive of intrahost infections or a persistent infection.

Dynamic behavior inside the SARS Cov2 genome suggests that the genome can be analyzed using a time series approach. Smoothed single nucleotide time series by rolling mean showed a series of oscillations in the content of the different nucleotides (Figure 15 A,C,E,G). Oscillations are nearly absent when the nucleotide frequency is grouped by the number of sunspots (Figure 15

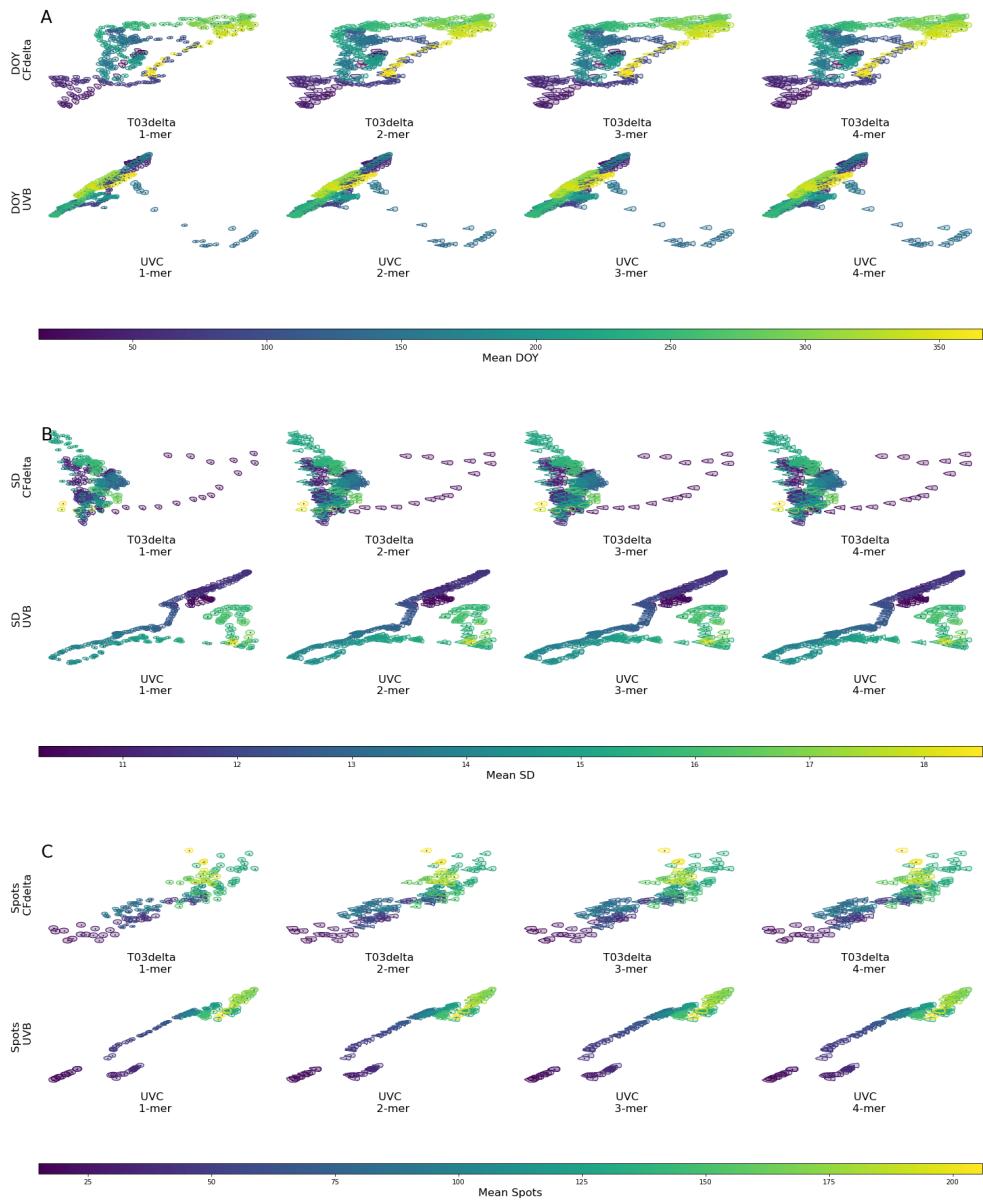


**Figure 15: Attractor-like behavior of SARS-CoV2 genome composition under different time scales.** A. Single nucleotide dynamics by date, vertical lines show the time on which different Remdesivir trials took place. Black and grey Wang et al and Beigel et al trials and red WHO consortium trial. B, D, F, and H Composition dynamics, color equal to the selected time scale, X and Y equal to [A] and [C] content, width and height of the marker are equal to [T] and [G] content. C Yearly single nucleotide dynamics. E Single nucleotide dynamics under a solar time scale, the number of sunspots are used as a measure of solar activity. G Yearly single nucleotide dynamics under the SD timescale.

337 E). Reconstructing the path followed by the four different nucleotides under different time scales  
 338 resulted in a series of attractor-like behavior when the data for the time scale contains one full cycle  
 339 (Figure 15 B,D,F,H). In the case of the SN time scale, only a single arc is observed.

340 Oscillatory behavior inside the SARS-CoV2 nucleotide composition can be used to schedule the  
 341 antiviral treatment with nucleotide-analogs. As changes in the nucleotide content will increase or  
 342 decrease the efficiency of those analogs. Particularly for SARS-CoV2 the different oscillations could  
 343 explain the mixed reports for Remdesivir an adenine analog at the beginning of the pandemic.  
 344 The earliest trial by Wang et al [33] was conducted from February to March the next one was  
 345 conducted by Beigel et al [34] from late February to middle April 2020 both trials found a benefit  
 346 from Remdesivir treatment. While the WHO solidarity consortium trial [35],which found little to no  
 347 effect of Remdesivir treatment, took place from late March to early October 2020 when the adenine  
 348 content was reduced several folds (Figure 15 A).

349 Although the previous observation holds better in a previous version of the analysis. In which  
 350 nucleotide content is evaluated for the complete genomic sequences as obtained from NCBI. Yet  
 351 errors or fragments added by the sequencing process itself could be biasing the previous observation.



**Figure 16: SARS-CoV2 genome adaptation under different time scales and environmental changes.** Genome composition is encoded as the shape of a closed curve, each column represents the k-mer composition and rows the time scale. Only highly correlated features were analyzed.

352 Current analysis focuses on the effective genome defined as the sequence between the first start  
 353 codon and the last stop codon. This most likely removes regulatory regions but it better represents  
 354 the SARS-CoV2 genome. Under the new analysis conditions, the reduction in adenine content is still  
 355 observed, however, the magnitude of the effect is less than previously presented. Nonetheless, early  
 356 trials were conducted at SD conditions with high adenine content while subsequent trials took place  
 357 at conditions of low adenine content. The magnitude of the effect might be miss represented but it

358 is still observed under different time scales.

359     The high dimensionality of the data hinders the ability to properly visualize how the genome  
360 adapts. One simple method to represent such information is by encoding the k-mer composition as  
361 the shape of a closed curve. This simple representation allows to analyze changes in composition  
362 under different time scales and their correlation to environmental variables previously selected.

363     K-mer composition under the environmental variables TO3delta and CFdelta grouped by SD  
364 showed a similar pattern as the one found in the stacked Kmer models. Similar shapes are near  
365 together and as the environmental variables change a transition between different shapes is also  
366 observed. While UVA and UVB showed a linear correlation between the environmental scale and  
367 viral genome composition and a smooth transition between the different shapes (Figure 16).

368     This new representation further confirms the dependence between SARS-Cov2 adaptation and  
369 solar activity. Also starts to provide a possible interpretation of the learned representation of the  
370 stacked K-mers models. The frequency of specific K-mers could encode an environmental advantage  
371 to the viral genome. As UV radiation is able to inactivate viruses by generating new bonds within  
372 paired bases it is possible that by reducing the frequency of specific K-mers such bonds are less likely  
373 to be formed [36]. Furthermore changes in genome size could also hint at an adaptation response to  
374 radiation. A smaller genome will be less likely to be inactivated by high levels of UV radiation that  
375 happen in the summer or at the Shwabe maximum. While a larger genome will be favored at low  
376 levels of radiation such as winter and the Shwabe minimum.

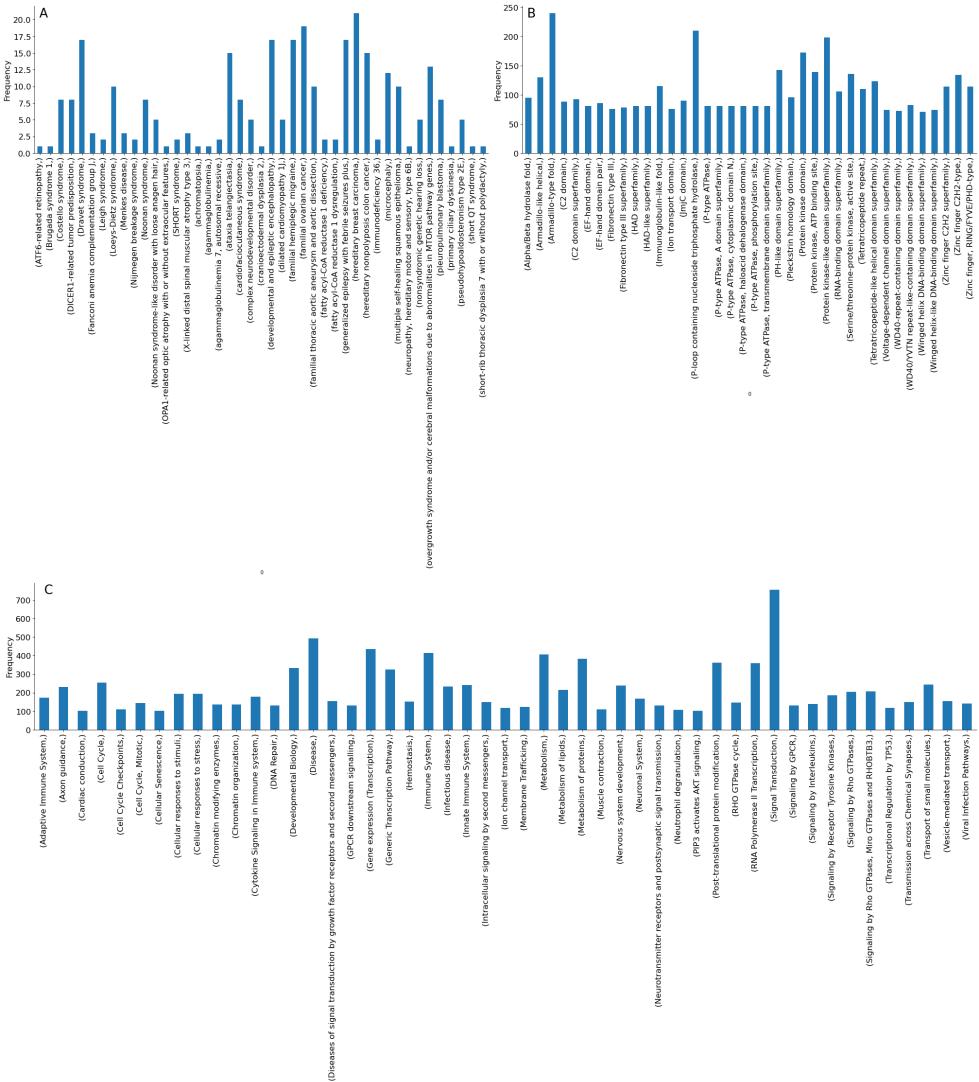
377     Changes in genome size could also affect the perceived severity of the infection by adding or  
378 removing parts of the genome. A larger genome will contain the full range of information to generate  
379 a wider range of symptoms. While a smaller one will be more constrained in that regard allowing  
380 the virus to infect with less perceived severity.

381     For example, Orf8 is a viral protein that plasma levels correlated with disease severity and  
382 mortality [37]. Infecting viruses lacking Orf8 was associated with better outcomes. Orf8 also inhibits  
383 the presentation of antigens by the MHC of class 1 [38]. The shrinking SARS-Cov2 genome could  
384 delete or partially delete this orf as an adaptation response to changing levels of radiation. Other  
385 Orfs inside the genome viruses could also be subjected to a similar adaptation mechanism.

### 386 **As resource optimization.**

387     Seasonal adaptation (genome rearrangement) and susceptibility(high spread rate) of the host to  
388 SARS-Cov2 point towards a specific dependence of SARS-Cov2 to cellular conditions. Specific  
389 seasonal expression patterns might drive the susceptibility to SARS-Cov2 as the main components  
390 of the virus are RNA and proteins. One possible mechanism used by the virus to adapt to those  
391 conditions will be to adapt its genome to the CUB of the host. CUB optimization effectively  
392 optimizes the nucleotide and amino acid availability inside the cell. Changes inside the genomic  
393 composition of the cell due to seasonal changes will alter the available resources to synthesize those  
394 genes, transcripts, and proteins. Thus even a virus highly adapted to the CUB of a particular cell  
395 type will face periods with low resources to replicate itself lowering its R0.

396     CUB is coded by the availability of specific codons inside the genes expressed inside a specific



**Figure 17: Data mining of compositionally similar transcripts.** A. Disease involvement of the selected transcripts. B. Most common domains coded by the selected transcripts. C. Pathway involvement of the proteins coded by the selected transcripts.

cell [39]. Seasonal gene expression will change a subset of genes leading to a dynamical CUB thus if a virus optimizing CUB will likely also follow seasonal gene expression. By mimicking the genetic composition of the cell a virus will be able to follow a particular seasonal pattern. Mimicry will also increase the chance of copying large enough pieces of a gene/transcript and synthesizing a viral protein with a fragment highly similar to the host leading to autoimmunity.

Seasonal genes with a similar composition to SARS-CoV2 will match both nucleotide and amino acid pools needed for the virus to replicate itself. To test that hypothesis a new dataset is generated with the GRCh 38 reference transcripts sequences and compared with the mean SARS-CoV2 composition using the Mahalanobis distance. Selected transcripts are then predicted with the trained

406 model and the transcripts with the lowest reconstruction error are selected. Information about the  
407 retrieved transcripts is then obtained with mygene service.

408 The previous filter results on a series of transcripts but information was obtained for only a  
409 subset of them. Interpro classification showed among the main coded domains were the armadillo-  
410 like domain, the Protein Kinase domain, and the zinc finger domain (Figure 17 B). Pathway analysis  
411 showed the involvement of the immune system, gene expression, and signal transduction among  
412 others (Figure 14 C). However, the involvement of GPCR signaling was observed under different  
413 categories. While the involvement of the different transcripts in diseases showed many hereditary  
414 disorders that affect the brain and the heart (Figure 17 A).

415 Although in the case of SARS-CoV2, the information regarding following seasonal expression  
416 patterns is scarce, mainly due to two reasons. The low amount of information regarding seasonal  
417 expression patterns. And that most of the functional information is skewed towards a few subsets  
418 of proteins/genes [40], however valuable information still can be retrieved.

419 For example, two seasonal genes CLOCK and RORA were selected by the previous filter. CLOCK  
420 is mainly expressed over the winter season, while RORA is expressed over the summer season [41].  
421 Direct involvement in COVID-19 has also been found for both genes. BMAL1 is a protein that  
422 interacts with CLOCK and controls the expression of the ACE2 receptor also regulating the entry  
423 of SARS-CoV2 [42]. While RORA is differentially expressed in COVID-19 as well as in heart failure  
424 [43]. These two examples show two mechanisms that can be concurrent. Differentially regulated  
425 genes could increase the susceptibility to COVID-19 like in the case of RORA. Genomic mimicry  
426 could copy a fragment of a host protein to facilitate the manipulation of the molecular machinery,  
427 as suggested by CLOCK. A third possibility is when a host protein presents self-interaction, in this  
428 case, the similarity is on a specific region. One example of this kind of mimicry is the TMPRSS15  
429 protein member of the TMPRSS and some members of this family undergo auto activation [44].  
430 TMPRSS15 is also found by the previous analysis and is another entry factor of SARS-CoV2 [45].

431 A manual search of the different genes results in several genes with little information regarding  
432 their direct involvement with SARS-CoV2 infection. Most of the genes didn't show any kind of  
433 involvement, however addressing them as a whole also gave a little insight into the possible subversion  
434 mechanism used by SARS-CoV2.

435 For example, the selection of different olfactory receptors, other GPCRs, channels, and trans-  
436 porters point towards the selection of a series of proteins that target the membrane. This specific  
437 targeting is important for the release of viral particles to the extracellular space [46]. Similarity with  
438 other proteins involved in endosomal trafficking such as Rab27B involved in the delivery of secretory  
439 granules [47]. Or Rab3C is involved in the last steps of exocytosis and the release of neurotransmit-  
440 ters [48]. HOOK1 another selected protein, plays a role in trafficking via the microtubule network  
441 [49]. Members of the SNX (SNX2, SNX16, SNX10, SNX14, SNX13) family regulate GPCR signaling  
442 by the trafficking of those receptors [50]. Those particular proteins contain a BAR domain that  
443 interacts with different membrane lipids in the membrane, this allows them to modify the shape of  
444 the membrane to create an endocytic vesicles [51]. Membrane composition can also be modified to  
445 create those anchoring points for BAR domains. Particularly PIP2 is regulated by proteins such as  
446 PI3K and INPP4 [52] and different isoforms of those proteins were found in the previous analysis.

447 Therefore the selection of highly similar transcripts could result in a partial characterization  
448 of the subverted signaling pathways. Mimmickyng eukaryotic motifs and domains are necessary to  
449 better manipulate the host molecular machinery. Mimicry can be optimized by copying or optimizing  
450 the CUB of a particular cell or tissue. Mutations then arise as a consequence of CUB optimization,  
451 as particular codons could be found in less quantity in the infected cell. Some mutations are silent as  
452 the mutation aims for a suitable synonym codon. While others change both the codon and the amino  
453 acid. Specifics on how mutations are selected or favored remain elusive, however, it is likely that the  
454 genome 3D structure encodes information regarding activation and mutation. The 3D structure of  
455 RNA can behave as a logic gate by hybridizing with different complementary fragments [53]. Host  
456 RNA fragments that result from Nsp15 digestion could hybridize with the complementary viral RNA  
457 linearizing the genome. Hybridization with digested mRNA fragments could provide the virus with a  
458 mechanism to address the molecular resources as well as to aid in viral genome replication. Allowing  
459 SARS-CoV2 to activate itself only at conditions where the resource and genomic landscape ease its  
460 replication.

461 Addressing the pathway involved by manually searching the function of the different selected  
462 genes continues to be challenging. Also as the virus continues to adapt to the environment, selected  
463 genes will likely change. This will skew the selection of different proteins within the same family and  
464 perform the same function. Therefore a broader selection could be more useful than a set of specific  
465 ones. However this only further confirms the need to develop an automated method to analyze the  
466 selected genes.

## 467 Post Acute Sequelae

468 Once the acute phase of the infection is cleared it is expected for the host to return to a normal  
469 state of health. However, in some cases, the acute infection can leave long-lasting health effects on  
470 the host. The lack of an established mechanism behind post-acute sequelae as well as the lack of  
471 molecular markers for the condition just makes the diagnosis more difficult. Referred to as Long  
472 COVID or PASC, a term that encapsulates the series of sequelae left by the acute infection. Although  
473 SARS-CoV2 is not the only virus that can leave long-term sequelae its fast spread as well as the  
474 heterogeneity of the symptoms make it a growing public health problem. Current PASC estimations  
475 show that around 10% of all SARS-CoV2 infections can lead to PASC[54].

476 The circadian nature of SARS-CoV2 points to circadian dysregulations in the post-acute phase.  
477 Particularly a light-based circadian dysregulation affecting in different degrees a series of phys-  
478 iological processes that rely on circadian rhythmicity. Specifics on how such dysregulations drive  
479 post-acute sequelae are presented as two different and non-excluding mechanisms. Nonetheless, both  
480 rely on reducing the amount of resources available for the virus to replicate itself. As such strategy  
481 has already been found experimentally [55]. This strategy by itself will not be able to prevent the  
482 spread of the virus, but it will lower the burden on the immune system.

483 **Epigenetic Manipulation.**

484 By each new subsequent infection the host is at risk of developing a persistent infection. In such  
485 scenario the infection is not cleared and different cells in the host remains infected. If the infection  
486 became persistent and affects other organs the host could try to lower the amount of resources  
487 and starve the virus. This blockade could be obtained by the downregulation of genes with similar  
488 composition to the infecting virus. Downregulation will also reduce the probability for the virus to  
489 copy a protein fragment that further refines the ability of the virus to manipulate cellular molecular  
490 machinery.

491 The previous mechanism proposes a one-sided response and frames the post-acute sequelae as  
492 a starving strategy. Yet the virus could also generate such dysregulation as a strategy to freeze  
493 in time those conditions that enable it to generate a large number of copies of itself. Thus viral  
494 infections could also generate epigenetic changes to increase the expression of compositional similar  
495 genes at a cellular level. This could create miscommunication between neighboring cells and prevent  
496 the function of the tissue as a whole.

497 Both mechanisms propose epigenetic changes that modulate resource availability and it could  
498 be complicated to assess the correctness of either of those. However, epigenetic changes occur post  
499 SARS-CoV2 infection [56], and manipulation of epigenetic machinery such as HDAC2 has already  
500 been reported [57]. Therefore epigenetic modifications are a result of SARS-CoV2 infection but its  
501 meaning remains to be addressed. Moreover, those repercussions might not be noticeable by the  
502 host until they reach a certain threshold.

503 Epigenetic modification could also increase the resources for another kind of virus, activating dor-  
504 mant infections. Viral activation after SARS-CoV2 infection has been previously reported, however,  
505 has not been linked to any kind of epigenetic modifications [58].

506 The suggested mechanisms rely on the dysregulation of genes with high compositional similarity  
507 to SARS-CoV2, thus identifying the phenotype that results from such dysregulation could provide  
508 insight into post-acute sequelae. Mentioning some of the previously selected genes CLOCK dysreg-  
509 ulation by gene mutations could result in delayed sleep phase disorder [59]. A disease that presents  
510 difficulties in sleep and misregulation of body temperature, symptoms already reported as post-acute  
511 sequelae of COVID-19 [60]. While RORA dysregulation by gene mutation could result in Intellectual  
512 Developmental Disorder with or Without Epilepsy or Cerebellar Ataxia, a condition with cognitive  
513 impairment [61], a symptom also reported for post-acute sequelae of COVID-19 [62].

514 Circadian dysregulation could also be the result of impairing other circadian signaling pathways.  
515 Particularly protein glycosylation prevents the degradation of proteins modifying their signaling  
516 time [63]. And calcium oscillatory signaling synchronizes the circadian rhythm in the central nervous  
517 system [64]. Lack of synchronization between the different oscillators, local and global could result  
518 in miscommunication between different systems inside the host. Glycosylation and calcium signaling  
519 could be influenced by genes selected from their high compositional similarity to SARS-CoV2 as well  
520 as to generate exercise intolerance by two different mechanisms.

521 AGL is a gene that encodes for the glycogen debranching enzyme, mutations in AGL can generate  
522 a condition known as Glycogen Storage Disease III. This particular disease is characterized by the

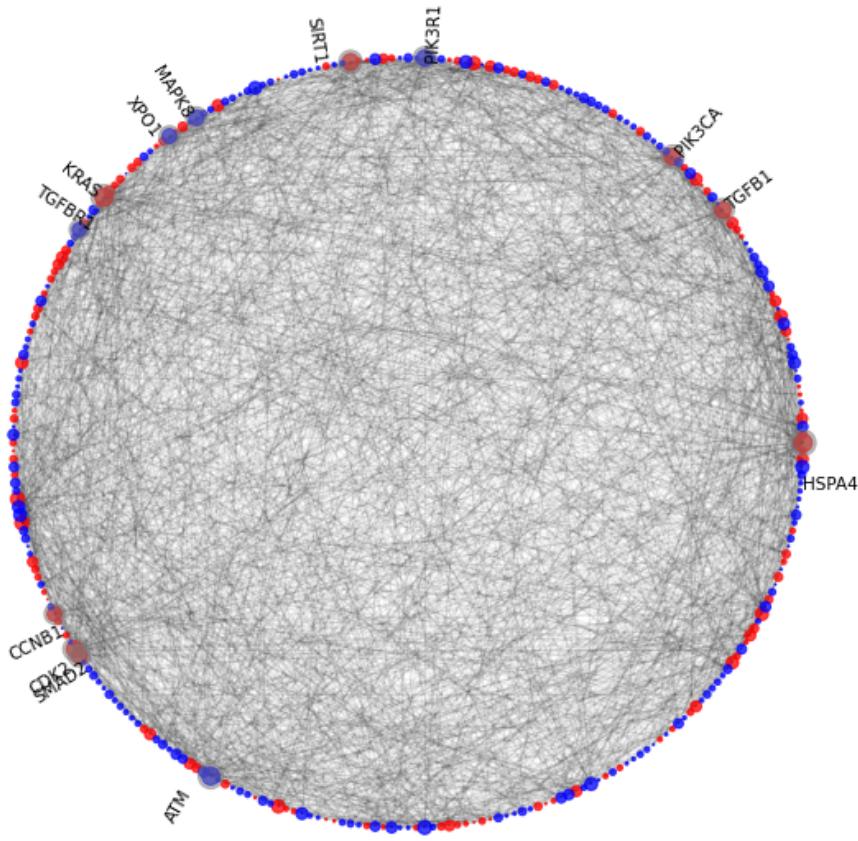


Figure 18: **Protein interactions between the selected compositionally similar genes.** Blue selected genes, red not selected.

523 accumulation of glycogen and affects the liver and the skeletal muscle [65]. This condition can  
 524 also generate exercise intolerance and the proposed mechanism behind is the result of low energy  
 525 availability to the skeletal muscle [66].

526 Calcium mishandling could also derive in exercise intolerance, particularly mutations in calmod-  
 527 ulin 2 lead to a condition known as catecholaminergic polymorphic ventricular tachycardia (CPVT).  
 528 Among the different symptoms of CPVT are abnormal heart rate and exercise intolerance. Mutations  
 529 inside calmodulin-2 that generate CPVT lower its affinity to calcium [67]. Calmodulin-2 is  
 530 also a compositionally similar gene to SARS cov2.

531 Framing post-acute sequela as the miss regulation of compositionally similar genes could help  
 532 to increase the number of possible treatments and will have the advantage to provide in some cases  
 533 already approved treatments. For example, lithium cation that targets CLOCK has been used to  
 534 increase the recovery rate in COVID-19 patients [68] as well as being evaluated as a treatment  
 for long Covid (ClinicalTrials.gov Identifier NCT05618587). Drugs targeting AGL also showed to

536 inhibit viral replication [69]. While treatments recommended for CPVT such as Flecainide showed  
537 a reduction in the odds of being hospitalized for COVID-19 [70]. Beta-blockers such as metropolol  
538 also showed a protective effect in COVID-19 patients [71], and it's being evaluated for long covid  
539 treatment (ClinicalTrials.gov Identifier NCT05096884).

540 Furthermore, if compositionally similar genes are dysregulated they could disrupt the function  
541 of a protein on which they converge due to their different interactions. Analyzing the interaction  
542 network of the compositionally similar genes showed a series of targets with a high number of  
543 interactions. Of all those different proteins only one presented the condition of interacting with a  
544 high number of compositionally similar genes and not being compositionally similar to SASRS-Cov2  
545 (Figure 18).

546 HSPA4 encodes for the HSP70 protein, a protein that has a neuroprotective role in many cerebral  
547 insults [72]. Its also involved in transient cerebral ischemia, a stroke that lasts for a few minutes,  
548 and the symptoms disappear between one and 24 hours. Symptoms include, numbness or weakness  
549 in the face, arm, or leg, especially on one side of the body, trouble seeing in one or both eyes,  
550 difficulty with walking, dizziness, confusion or difficulty in talking or understanding speech, loss of  
551 balance and coordination. HSP70 is also manipulated by SARS-Cov2 infection making it another  
552 likely target for treatment development [73].

## 553 Nucleotide Starvation.

554 Another mechanism by which the host could starve the virus is by simply reducing the amount  
555 of specific nucleotides. As shown before each nucleotide have periods of low and high demand,  
556 lowering the amount of a specific nucleotide will starve the virus in the long term. However the  
557 inverse problem could also be true, and the virus attempts to increase the amount of the limiting  
558 factor. In both cases, the nucleotide supply will be disrupted by either a lack of supply or a relocation  
559 of the supply. Specific kind of nucleotide will also leave an imprint in the kind of post-acute sequelae  
560 generated in the host.

561 Adenine is one of the four different nucleotides needed to make the SARS-Cov2 genetic material.  
562 Adenine, and other chemical derivatives of adenine, have also a particular role in the circadian  
563 rhythm. As adenosine levels build-up signalling the sleep/wake cycle through the adenosine receptor  
564 [74]. Adenine is also the building block of the second messenger cAMP, this particular molecule has  
565 a role in a series of biological processes [75]. Orthostasis (going from a lay position to an upright  
566 position) is followed by an increase in plasma cAMP [76]. Intracellular cAMP can be generated  
567 by GPCR signaling or by calcium. Particularly in myocytes intracellular cAMP is secreted to the  
568 extracellular space and converted in adenosine to then be transported into the intracellular space.  
569 This movement of different chemical derivates of adenine generates a series of oscillations that aid to  
570 dampen beta-adrenergic receptors signaling [77]. This regulatory effect could also help to modulate  
571 calcium signaling in myocytes. [78] Another oscillatory mechanism that involves cAMP and calcium  
572 signalling has been found in mouse beta cells [79]. If the adenine supply is disrupted then all  
573 those different mechanisms could be impacted at different degrees. For example, low adenine levels  
574 could hinder the total cAMP. This will lower the amount of cAMP to be converted into adenosine,

575 without a dampening signal myocyte contractions goes without check leading to abnormal heart  
576 rhythm. Thus in conditions that lead to an increase in cAMP such as orthostasis, this mechanism  
577 will be impaired and may lead to postural orthostatic tachycardia POTS.

578 The possible involvement of other nucleotides remains to be addressed, however, the impairment  
579 of nucleotide pools could also have other repercussions. Disruption of nucleotide pools particu-  
580 larly specific rNTPs can disrupt mitochondrial function. Particularly disruption of pyrimidine supply  
581 into mitochondria lowers mitochondrial gene expression. Leading to the loss of components of the  
582 OXPHOS system, lowering mitochondrial respiration and the production of ATP. [80]

583 Although previous mechanisms just try to explain a subset of the different symptoms experienced  
584 by long covid patients. They provide a series of possible targets for either treatment development or  
585 possible molecular markers. Epigenetic modifications or changes in gene expression patterns might  
586 be harder to address in the general population. But changes in nucleotides and nucleotide derivatives  
587 might be easier to adopt at a large scale.

588 Also, the involvement of the immune system is largely overlooked as is almost impossible to  
589 address it in silico. Nonetheless by looking at different mechanisms that are independent of the  
590 immune system a previously disregarded part of the phenomenon is being presented. Combining  
591 both approaches most likely will provide a better understanding of long COVID. as viral persistence  
592 might be the driving force behind such mechanisms.

593 The specific variant, infection date, and previous infections most likely will determine the spe-  
594 cific kind of dysregulation. Viral persistence will also increase the risk of copying fragments of host  
595 proteins and developing autoimmunity as well as immune activation. Thus addressing viral persis-  
596 tence will need to be concomitant with other treatments aiming to restore or manage some of the  
597 symptoms of post-acute sequelae.

598 As more than one intervention will likely be needed to restore post-acute sequelae there's a higher  
599 risk of interaction between treatments. Thus careful considerations need to be taken to design an  
600 adjust possible treatments.

## 601 Conclusions

- 602 • The existence of different molecular clocks encoded inside the SARS-Cov2 genomic sequence  
603 shows a deterministic mechanism used by the virus to replicate its genome. Other clock  
604 components remain to be found if they exist. Nevertheless, the suggested random nature of  
605 viral mutation is less than previously suggested or non-existing.
- 606 • Detection of other frequency components inside solar activity and radiation could provide other  
607 periods of high SARS-Cov2 susceptibility.
- 608 • Non-linearity inside the environmental driver for SARS-Cov2 adaptation and susceptibility  
609 adds a confounding variable to the different risk/benefit assessments. Interpretation of the  
610 efficacy and safety of pharmacological and non-pharmacological interventions might need a  
611 reanalysis.

- 612     ● SD is not a perfect time scale for SARS-Cov2 seasonality but it offers an easy metric to follow  
613       for the implementation of different risk reduction measures as well as vaccination schedules.
- 614     ● Addressing nutritional deficiencies that impact circadian signaling could be used as a prophylactic  
615       treatment at periods of low solar activity for protection against pathogens with circadian  
616       seasonality.
- 617     ● Low-frequency components of low solar activity provide an easy metric for pandemic preparedness.
- 619     ● Although SARS-Cov2 follows solar radiation other consequences of low solar activity could  
620       also ease its transmission.
- 621     ● Targeting low-variable components inside the SARS Cov2 genome might provide a better  
622       long-term treatment strategy.
- 623     ● Environmental correlation to genomic composition can be used for the scheduling of possible  
624       treatments.
- 625     ● Changes in variation patterns can be used to skew treatments towards an immunization based  
626       or an antiviral based.
- 627     ● Further improvements in the different sequence representation schemes and machine learning  
628       models will fast-track the characterization of emerging and neglected pathogens.

## 629     **Acknowledgments**

630     Access to data sources from different government organizations was key for the development of the  
631       previous analysis. I would like to thank everyone involved in the acquisition, curation, and other  
632       support staff as well as the policymakers who made the data available to anyone. Publicly available  
633       data offers a model of silent collaboration from which everyone benefits from it.

## 634     **Declarations**

635     The author declares no competing interests. Although the author is trying to raise funds, with little  
636       to unexisting success :’3, to continue research under different platforms listed here. Linktree

## 637     **Methods**

638     Genomic sequences were obtained from the NCBI SARS-Cov2 resource site up to june 2023 [81].  
639     Only complete genomes without any ambiguous characters were selected for the analysis. From the  
640       selected genomes the effective genome is extracted by selecting the fragment between the first start  
641       codon and the last stop codon. This selection is done regardless of the reading frame in which the  
642       start and stop points are found, and also regardless of the kind of stop codon. Available metadata

643 was used to reverse geolocate the corresponding geographical location. The isolation date was used  
644 as a time descriptor for each sequence as well as for the calculation of the SD [82].

645 Three sequence representations schemes are proposed to analyze SARS Cov 2 genomes. The first  
646 one consists of the stacked frequency of k-size fragments or k-mers. SARS-CoV2 genome sequences  
647 are spitted in a sliding manner leading to fragments with an overlap of (k-1) nucleotides and (n-k)  
648 fragments. Frequency is normalized for each k and the different vectors are combined in a single  
649 one. The same method is applied to the reference transcripts dataset GRCh 38.

650 For the second one, each sequence is divided into 16 fragments of equal size. Then for each  
651 fragment, a graph is constructed by setting as nodes the 2-mer and adding a link between consecutive  
652 2-mers derived from a sliding fragmentation. The same process is applied to the reversed fragment  
653 and the normalized adjacency matrix is used as sequence representation.

654 The third dataset consists of the one-hot-encoded full genomic sequence, yet the sequence is  
655 transformed into a 4D array. Where the first three dimensions represent the location of each base  
656 in the genome and the last one is the kind of nucleotide. This specific rearrangement could bring  
657 closer different genomic sequences, mimicking the 3D arrangement of the viral genome inside the  
658 viral particle.

659 For each dataset, a specific machine learning model is used in a representation learning task. For  
660 the stacked k-mers dataset the VAE architecture consists of a multilayer perceptron, and a modified  
661 version to retrieve a dynamical system. For the remaining datasets, the VAE architecture consists of  
662 a deep convolutional network of different architectures for each dataset. A five-fold cross-validation  
663 was applied and only the best fold is shown.

664 Case data were obtained by the different government agencies involved, as testing and update  
665 frequency changed between countries data was collected up to May 2022 [83] [84] [85] [86] [87] [88]  
666 [89] [90] [91]. Reverse geolocation is used to retrieve the latitude and longitude at a city level for  
667 the available epidemic curves.

668 Solar activity data consists of sunspots data from the WDC-SILSO Royal Observatory of Belgium  
669 Brussels [92] and the TSIS SIM level 3 solar spectral irradiance 24hr means [93]. While environmental  
670 features are obtained from the Aqua/AIRS L3 Daily Standard Physical Retrieval dataset. Missing  
671 values are predicted with a random forest regressor trained on a sliding window of 70 days [94]. A  
672 complete summary of the different scripts and datasets can be found at Github

## 673 References

- 674 [1] Á. O'Toole *et al.*, "Assignment of epidemiological lineages in an emerging pandemic using  
675 the pangolin tool," *Virus Evolution*, vol. 7, no. 2, Jul. 2021, veab064, ISSN: 2057-1577. DOI:  
676 10.1093/ve/veab064. eprint: <https://academic.oup.com/ve/article-pdf/7/2/veab064/49179570/veab064.pdf>. [Online]. Available: <https://doi.org/10.1093/ve/veab064>.
- 677 [2] D. Bank, N. Koenigstein, and R. Giryes, *Autoencoders*, 2021. arXiv: 2003.05991 [cs.LG].

- 679 [3] T. Trinh, A. Dai, T. Luong, and Q. Le, “Learning longer-term dependencies in RNNs with  
680 auxiliary losses,” in *Proceedings of the 35th International Conference on Machine Learning*, J.  
681 Dy and A. Krause, Eds., ser. Proceedings of Machine Learning Research, vol. 80, PMLR, Jul.  
682 2018, pp. 4965–4974. [Online]. Available: <https://proceedings.mlr.press/v80/trinh18a.html>.
- 683
- 684 [4] M. Afsar *et al.*, “Drug targeting nsp1-ribosomal complex shows antiviral activity against sars-  
685 cov-2,” *eLife*, vol. 11, S. Haider, M. Zaidi, and S. Haider, Eds., e74877, Mar. 2022, ISSN:  
686 2050-084X. DOI: 10.7554/eLife.74877. [Online]. Available: <https://doi.org/10.7554/eLife.74877>.
- 687
- 688 [5] H.-T. Kao, A. Orry, M. G. Palfreyman, and B. Porton, “Synergistic interactions of repurposed  
689 drugs that inhibit nsp1, a major virulence factor for covid-19,” *Scientific Reports*, vol. 12,  
690 no. 1, p. 10174, Jun. 2022, ISSN: 2045-2322. DOI: 10.1038/s41598-022-14194-x. [Online].  
691 Available: <https://doi.org/10.1038/s41598-022-14194-x>.
- 692
- 693 [6] B. Oronsky *et al.*, “Nucleocapsid as a next-generation covid-19 vaccine candidate,” *International  
694 Journal of Infectious Diseases*, vol. 122, pp. 529–530, Sep. 2022, ISSN: 1201-9712. DOI:  
695 10.1016/j.ijid.2022.06.046. [Online]. Available: <https://doi.org/10.1016/j.ijid.2022.06.046>.
- 696
- 697 [7] M. S. Khan *et al.*, “Adenovirus-vectored SARS-CoV-2 vaccine expressing S1-N fusion protein,”  
698 *Antibody Therapeutics*, vol. 5, no. 3, pp. 177–191, Jul. 2022, ISSN: 2516-4236. DOI: 10.1093/  
699 abt/tbac015. eprint: <https://academic.oup.com/abt/article-pdf/5/3/177/45348595/tbac015.pdf>. [Online]. Available: <https://doi.org/10.1093/abt/tbac015>.
- 700
- 701 [8] E. K. V. B. Silva *et al.*, “Immunization with sars-cov-2 nucleocapsid protein triggers a pul-  
702 monary immune response in rats,” *PLOS ONE*, vol. 17, no. 5, pp. 1–17, May 2022. DOI:  
703 10.1371/journal.pone.0268434. [Online]. Available: <https://doi.org/10.1371/journal.pone.0268434>.
- 704
- 705 [9] K. L. O’Donnell, T. Gourdine, P. Fletcher, C. S. Clancy, and A. Marzi, “Protection from  
706 covid-19 with a vsv-based vaccine expressing the spike and nucleocapsid proteins,” *Frontiers  
707 in Immunology*, vol. 13, 2022, ISSN: 1664-3224. DOI: 10.3389/fimmu.2022.1025500. [Online].  
708 Available: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.1025500>.
- 709
- 710 [10] J. Martel, S.-H. Chang, G. Chevalier, D. M. Ojcius, and J. D. Young, “Influence of electromag-  
711 netic fields on the circadian rhythm: Implications for human health and disease,” *Biomedical  
712 Journal*, vol. 46, no. 1, pp. 48–59, 2023, ISSN: 2319-4170. DOI: <https://doi.org/10.1016/j.bj.2023.01.003>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2319417023000033>.
- 713
- 714 [11] T. A. G. on SARS-CoV-2 Virus Evolution, “Historical working definitions and primary ac-  
715 tions for sars-cov-2 variants,” Mar. 2023. [Online]. Available: <https://www.who.int/publications/m/item/historical-working-definitions-and-primary-actions-for-sars-cov-2-variants>.
- 716

- 717 [12] S. W. P. C. ( U.S. Dept. of Commerce NOAA, “Solar cycle progression,” Apr. 2023. [Online].  
718 Available: <https://www.swpc.noaa.gov/products/solar-cycle-progression>.
- 719 [13] A. E. Samoilov *et al.*, “Case report: Change of dominant strain during dual sars-cov-2 in-  
720 fection,” *BMC Infectious Diseases*, vol. 21, no. 1, p. 959, Sep. 2021, ISSN: 1471-2334. DOI:  
721 [10.1186/s12879-021-06664-w](https://doi.org/10.1186/s12879-021-06664-w). [Online]. Available: <https://doi.org/10.1186/s12879-021-06664-w>.
- 723 [14] H.-Y. Zhou *et al.*, “Genomic evidence for divergent co-infections of sars-cov-2 lineages,” *bioRxiv*,  
724 2021. DOI: [10.1101/2021.09.03.458951](https://doi.org/10.1101/2021.09.03.458951). eprint: <https://www.biorxiv.org/content/early/2021/09/04/2021.09.03.458951.full.pdf>. [Online]. Available: <https://www.biorxiv.org/content/early/2021/09/04/2021.09.03.458951>.
- 727 [15] T. Karamitros, G. Papadopoulou, M. Bousali, A. Mexias, S. Tsiodras, and A. Mentis, “Sars-  
728 cov-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies,”  
729 *Journal of Clinical Virology*, vol. 131, p. 104585, 2020, ISSN: 1386-6532. DOI: <https://doi.org/10.1016/j.jcv.2020.104585>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1386653220303279>.
- 732 [16] F. Chen *et al.*, “Dissimilation of synonymous codon usage bias in virus–host coevolution due  
733 to translational selection,” *Nature Ecology & Evolution*, vol. 4, no. 4, pp. 589–600, Apr. 2020,  
734 ISSN: 2397-334X. DOI: [10.1038/s41559-020-1124-7](https://doi.org/10.1038/s41559-020-1124-7). [Online]. Available: <https://doi.org/10.1038/s41559-020-1124-7>.
- 736 [17] J. B. Gibbons *et al.*, “Association between vitamin d supplementation and covid-19 infection  
737 and mortality,” *Scientific Reports*, vol. 12, no. 1, p. 19397, Nov. 2022, ISSN: 2045-2322. DOI:  
738 [10.1038/s41598-022-24053-4](https://doi.org/10.1038/s41598-022-24053-4). [Online]. Available: <https://doi.org/10.1038/s41598-022-24053-4>.
- 740 [18] T. I. Hariyanto, D. Intan, J. E. Hananto, H. Harapan, and A. Kurniawan, “Vitamin d supple-  
741 mentation and covid-19 outcomes: A systematic review, meta-analysis and meta-regression,”  
742 *Reviews in Medical Virology*, vol. 32, no. 2, e2269, 2022. DOI: <https://doi.org/10.1002/rmv.2269>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/rmv.2269>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/rmv.2269>.
- 745 [19] S.-H. Lan, H.-Z. Lee, C.-M. Chao, S.-P. Chang, L.-C. Lu, and C.-C. Lai, “Efficacy of mel-  
746 atonin in the treatment of patients with covid-19: A systematic review and meta-analysis of  
747 randomized controlled trials,” *Journal of Medical Virology*, vol. 94, no. 5, pp. 2102–2107, 2022.  
748 DOI: <https://doi.org/10.1002/jmv.27595>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmv.27595>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.27595>.
- 751 [20] P. C. Pereira, C. J. de Lima, A. B. Fernandes, R. A. Zângaro, and A. B. Villaverde, “Car-  
752 diopulmonary and hematological effects of infrared led photobiomodulation in the treatment of  
753 sars-cov2,” *Journal of Photochemistry and Photobiology B: Biology*, vol. 238, p. 112619, 2023,  
754 ISSN: 1011-1344. DOI: <https://doi.org/10.1016/j.jphotobiol.2022.112619>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1011134422002342>.

- 756 [21] M. Zvyagin *et al.*, “Genslms: Genome-scale language models reveal sars-cov-2 evolutionary  
757 dynamics,” *bioRxiv*, 2022. DOI: 10.1101/2022.10.10.511571. eprint: <https://www.biorxiv.org/content/early/2022/11/23/2022.10.10.511571.full.pdf>. [Online]. Available:  
758 <https://www.biorxiv.org/content/early/2022/11/23/2022.10.10.511571>.
- 760 [22] D. Lebatteux, H. Soudeyns, I. Boucoiran, S. Gantt, and A. B. Diallo, “Machine learning-  
761 based approach kevolve efficiently identifies sars-cov-2 variant-specific genomic signatures,”  
762 *bioRxiv*, 2022. DOI: 10.1101/2022.02.07.479343. eprint: <https://www.biorxiv.org/content/early/2022/02/07/2022.02.07.479343.full.pdf>. [Online]. Available: <https://www.biorxiv.org/content/early/2022/02/07/2022.02.07.479343>.
- 765 [23] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, “Data-driven discovery of coordinates  
766 and governing equations,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 45,  
767 pp. 22 445–22 451, 2019. DOI: 10.1073/pnas.1906995116. eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.1906995116>. [Online]. Available: <https://www.pnas.org/doi/abs/10.1073/pnas.1906995116>.
- 770 [24] B. Ross, “The development of fractional calculus 1695–1900,” *Historia Mathematica*, vol. 4,  
771 no. 1, pp. 75–89, 1977, ISSN: 0315-0860. DOI: [https://doi.org/10.1016/0315-0860\(77\)90039-8](https://doi.org/10.1016/0315-0860(77)90039-8). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0315086077900398>.
- 774 [25] J. Calbó, D. Pagès, and J.-A. González, “Empirical studies of cloud effects on uv radiation:  
775 A review,” *Reviews of Geophysics*, vol. 43, no. 2, 2005. DOI: <https://doi.org/10.1029/2004RG000155>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2004RG000155>. [Online]. Available: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004RG000155>.
- 779 [26] C. D. McNaughton, N. M. Adams, C. H. Johnson, M. J. Ward, J. E. Schmitz, and T. A.  
780 Lasko, “Diurnal variation in sars-cov-2 pcr test results: Test accuracy may vary by time of  
781 day,” *Journal of Biological Rhythms*, vol. 36, no. 6, pp. 595–601, 2021, PMID: 34696614. DOI:  
782 10.1177/07487304211051841. eprint: <https://doi.org/10.1177/07487304211051841>.  
783 [Online]. Available: <https://doi.org/10.1177/07487304211051841>.
- 784 [27] A. V. Winnett *et al.*, “Morning sars-cov-2 testing yields better detection of infection due to  
785 higher viral loads in saliva and nasal swabs upon waking,” *Microbiology Spectrum*, vol. 10,  
786 no. 6, e03873–22, 2022. DOI: 10.1128/spectrum.03873-22. eprint: <https://journals.asm.org/doi/pdf/10.1128/spectrum.03873-22>. [Online]. Available: <https://journals.asm.org/doi/abs/10.1128/spectrum.03873-22>.
- 789 [28] X. Zhuang *et al.*, “Time-of-day variation in sars-cov-2 rna levels during the second wave of  
790 covid-19,” *Viruses*, vol. 14, no. 8, 2022, ISSN: 1999-4915. DOI: 10.3390/v14081728. [Online].  
791 Available: <https://www.mdpi.com/1999-4915/14/8/1728>.

- 792 [29] M. Romdhani *et al.*, “Is there a diurnal variation of covid-19 patients warranting presentation  
793 to the health centre? a chronobiological observational cross-sectional study,” *Annals of Medicine*,  
794 vol. 54, no. 1, pp. 3059–3067, 2022, PMID: 36308396. DOI: 10.1080/07853890.2022.2136399.  
795 eprint: <https://doi.org/10.1080/07853890.2022.2136399>. [Online]. Available:  
796 <https://doi.org/10.1080/07853890.2022.2136399>.
- 797 [30] J. Feynman and A. Ruzmaikin, “The centennial gleissberg cycle and its association with ex-  
798 tended minima,” *Journal of Geophysical Research: Space Physics*, vol. 119, no. 8, pp. 6027–  
799 6041, 2014. DOI: <https://doi.org/10.1002/2013JA019478>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JA019478>. [Online]. Available:  
800 <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JA019478>.
- 802 [31] C. B. Jackson, M. Farzan, B. Chen, and H. Choe, “Mechanisms of sars-cov-2 entry into cells,”  
803 *Nature Reviews Molecular Cell Biology*, vol. 23, no. 1, pp. 3–20, Jan. 2022, ISSN: 1471-0080. DOI:  
804 10.1038/s41580-021-00418-x. [Online]. Available: <https://doi.org/10.1038/s41580-021-00418-x>.
- 806 [32] C. McAloon *et al.*, “Incubation period of covid-19: A rapid systematic review and meta-  
807 analysis of observational research,” *BMJ Open*, vol. 10, no. 8, 2020, ISSN: 2044-6055. DOI:  
808 10.1136/bmjopen-2020-039652. eprint: <https://bmjopen.bmjjournals.com/content/10/8/e039652.full.pdf>. [Online]. Available:  
809 <https://bmjopen.bmjjournals.com/content/10/8/e039652>.
- 810 [33] Y. Wang *et al.*, “Remdesivir in adults with severe covid-19: A randomised, double-blind,  
811 placebo-controlled, multicentre trial,” *The Lancet*, vol. 395, no. 10236, pp. 1569–1578, 2020,  
812 ISSN: 0140-6736. DOI: [https://doi.org/10.1016/S0140-6736\(20\)31022-9](https://doi.org/10.1016/S0140-6736(20)31022-9). [Online]. Available:  
813 <https://www.sciencedirect.com/science/article/pii/S0140673620310229>.
- 814 [34] J. H. Beigel *et al.*, “Remdesivir for the treatment of covid-19 — final report,” *New England  
815 Journal of Medicine*, vol. 383, no. 19, pp. 1813–1826, 2020, PMID: 32445440. DOI: 10.1056/  
816 NEJMoa2007764. eprint: <https://doi.org/10.1056/NEJMoa2007764>. [Online]. Available:  
817 <https://doi.org/10.1056/NEJMoa2007764>.
- 818 [35] “Repurposed antiviral drugs for covid-19 — interim who solidarity trial results,” *New England  
819 Journal of Medicine*, vol. 384, no. 6, pp. 497–511, 2021, PMID: 33264556. DOI: 10.1056/  
820 NEJMoa2023184. eprint: <https://doi.org/10.1056/NEJMoa2023184>. [Online]. Available:  
821 <https://doi.org/10.1056/NEJMoa2023184>.
- 822 [36] R. Saha and I. A. Chen, “Effect of uv radiation on fluorescent rna aptamers’ functional and  
823 templating ability,” *ChemBioChem*, vol. 20, no. 20, pp. 2609–2617, 2019. DOI: <https://doi.org/10.1002/cbic.201900261>. eprint: <https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/cbic.201900261>. [Online]. Available:  
824 <https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/cbic.201900261>.
- 827 [37] X. Wu *et al.*, “Secreted orf8 is a pathogenic cause of severe covid-19 and is potentially targetable  
828 with select nlrp3 inhibitors,” *bioRxiv*, 2022. DOI: 10.1101/2021.12.02.470978. eprint: <https://www.biorxiv.org/content/early/2022/08/22/2021.12.02.470978.full.pdf>. [Online]. Available:  
829 <https://www.biorxiv.org/content/early/2022/08/22/2021.12.02.470978>.

- 831 [38] L. Zinzula, "Lost in deletion: The enigmatic orf8 protein of sars-cov-2," *Biochemical and Bio-*  
832 *physical Research Communications*, vol. 538, pp. 116–124, 2021, COVID-19, ISSN: 0006-291X.  
833 DOI: <https://doi.org/10.1016/j.bbrc.2020.10.045>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0006291X20319628>.
- 835 [39] S. T. Parvathy, V. Udayasuriyan, and V. Bhadana, "Codon usage bias," *Molecular Biology*  
836 *Reports*, vol. 49, no. 1, pp. 539–565, Jan. 2022, ISSN: 1573-4978. DOI: 10.1007/s11033-021-  
837 06749-4. [Online]. Available: <https://doi.org/10.1007/s11033-021-06749-4>.
- 838 [40] E. Dolgin, "The most popular genes in the human genome," *Nature*, vol. 551, no. 7681, pp. 427–  
839 431, Nov. 2017, ISSN: 1476-4687. DOI: 10.1038/d41586-017-07291-9. [Online]. Available:  
840 <https://doi.org/10.1038/d41586-017-07291-9>.
- 841 [41] X. C. Dopico *et al.*, "Widespread seasonal gene expression reveals annual differences in human  
842 immunity and physiology," *Nature Communications*, vol. 6, no. 1, p. 7000, May 2015, ISSN:  
843 2041-1723. DOI: 10.1038/ncomms8000. [Online]. Available: <https://doi.org/10.1038/ncomms8000>.
- 845 [42] X. Zhuang *et al.*, "The circadian clock component bmal1 regulates sars-cov-2 entry and repli-  
846 cation in lung epithelial cells," *iScience*, vol. 24, no. 10, p. 103144, 2021, ISSN: 2589-0042.  
847 DOI: <https://doi.org/10.1016/j.isci.2021.103144>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2589004221011123>.
- 849 [43] G. Zhang *et al.*, "Uncovering the genetic links of sars-cov-2 infections on heart failure co-  
850 morbidity by a systems biology approach," *ESC Heart Failure*, vol. 9, no. 5, pp. 2937–2954,  
851 2022. DOI: <https://doi.org/10.1002/ehf2.14003>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ehf2.14003>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ehf2.14003>.
- 854 [44] T. H. Bugge, T. M. Antalis, and Q. Wu, "Type ii transmembrane serine proteases \*jsupzj/supz,"  
855 *Journal of Biological Chemistry*, vol. 284, no. 35, pp. 23177–23181, Aug. 2009, ISSN: 0021-9258.  
856 DOI: 10.1074/jbc.R109.021006. [Online]. Available: <https://doi.org/10.1074/jbc.R109.021006>.
- 858 [45] Q. Zhang *et al.*, "Environmentally-induced mdig is a major contributor to the severity of  
859 covid-19 through fostering expression of sars-cov-2 receptor nrps and glycan metabolism,"  
860 *bioRxiv*, 2021. DOI: 10.1101/2021.01.31.429010. eprint: <https://www.biorxiv.org/content/early/2021/02/01/2021.01.31.429010.full.pdf>. [Online]. Available: <https://www.biorxiv.org/content/early/2021/02/01/2021.01.31.429010>.
- 863 [46] P. V'kovski, A. Kratzel, S. Steiner, H. Stalder, and V. Thiel, "Coronavirus biology and repli-  
864 cation: Implications for sars-cov-2," *Nature Reviews Microbiology*, vol. 19, no. 3, pp. 155–  
865 170, Mar. 2021, ISSN: 1740-1534. DOI: 10.1038/s41579-020-00468-6. [Online]. Available:  
866 <https://doi.org/10.1038/s41579-020-00468-6>.

- 867 [47] H. Gomi, K. Mori, S. Itohara, and T. Izumi, “Rab27b is expressed in a wide range of exocytic  
868 cells and involved in the delivery of secretory granules near the plasma membrane,” *Molecular*  
869 *Biology of the Cell*, vol. 18, no. 11, pp. 4377–4386, 2007, PMID: 17761531. DOI: 10.1091/mbe.  
870 e07-05-0409. eprint: <https://doi.org/10.1091/mbe.e07-05-0409>. [Online]. Available:  
871 <https://doi.org/10.1091/mbe.e07-05-040>.
- 872 [48] G. Fischer von Mollard, B. Stahl, A. Khokhlatchev, T. Südhof, and R. Jahn, “Rab3c is a  
873 synaptic vesicle protein that dissociates from synaptic vesicles after stimulation of exocytosis,”  
874 *Journal of Biological Chemistry*, vol. 269, no. 15, pp. 10971–10974, 1994, ISSN: 0021-9258.  
875 DOI: [https://doi.org/10.1016/S0021-9258\(19\)78076-4](https://doi.org/10.1016/S0021-9258(19)78076-4). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0021925819780764>.
- 877 [49] M. A. Olenick, R. Dominguez, and E. L. Holzbaur, “Dynein activator Hook1 is required for  
878 trafficking of BDNF-signaling endosomes in neurons,” *Journal of Cell Biology*, vol. 218, no. 1,  
879 pp. 220–233, Oct. 2018, ISSN: 0021-9525. DOI: 10.1083/jcb.201805016. eprint: [https://rupress.org/jcb/article-pdf/218/1/220/1379611/jcb\\_201805016.pdf](https://rupress.org/jcb/article-pdf/218/1/220/1379611/jcb_201805016.pdf). [Online].  
880 Available: <https://doi.org/10.1083/jcb.201805016>.
- 882 [50] B. Amatya *et al.*, “Snx-pxa-rgs-pxc subfamily of snxs in the regulation of receptor-mediated  
883 signaling and membrane trafficking,” *International Journal of Molecular Sciences*, vol. 22,  
884 no. 5, 2021, ISSN: 1422-0067. DOI: 10.3390/ijms22052319. [Online]. Available: <https://www.mdpi.com/1422-0067/22/5/2319>.
- 886 [51] S. Ahmed, W. Bu, R. T. C. Lee, S. Maurer-Stroh, and W. I. Goh, “F-bar domain proteins,”  
887 *Communicative & Integrative Biology*, vol. 3, no. 2, pp. 116–121, 2010, PMID: 20585502. DOI:  
888 10.4161/cib.3.2.10808. eprint: <https://doi.org/10.4161/cib.3.2.10808>. [Online].  
889 Available: <https://doi.org/10.4161/cib.3.2.10808>.
- 890 [52] H. Wang *et al.*, “Phosphatidylinositol 3,4-bisphosphate synthesis and turnover are spatially seg-  
891 regated in the endocytic pathway,” *Journal of Biological Chemistry*, vol. 295, no. 4, pp. 1091–  
892 1104, 2020, ISSN: 0021-9258. DOI: [https://doi.org/10.1016/S0021-9258\(17\)49918-2](https://doi.org/10.1016/S0021-9258(17)49918-2). [On-  
893 line]. Available: <https://www.sciencedirect.com/science/article/pii/S0021925817499182>.
- 894 [53] A. A. Green, J. Kim, D. Ma, P. A. Silver, J. J. Collins, and P. Yin, “Complex cellular logic  
895 computation using ribocomputing devices,” *Nature*, vol. 548, no. 7665, pp. 117–121, Aug.  
896 2017, ISSN: 1476-4687. DOI: 10.1038/nature23271. [Online]. Available: <https://doi.org/10.1038/nature23271>.
- 898 [54] T. Thaweethai *et al.*, “Development of a definition of postacute sequelae of sars-cov-2 infec-  
899 tion,” *JAMA*, vol. 329, no. 22, pp. 1934–1946, Jun. 2023, ISSN: 0098-7484. DOI: 10.1001/jama.  
900 2023.8823. eprint: [https://jamanetwork.com/journals/jama/articlepdf/2805540/jama\\_thaweethai\\_2023\\_oi\\_230062\\_1686240966.39381.pdf](https://jamanetwork.com/journals/jama/articlepdf/2805540/jama_thaweethai_2023_oi_230062_1686240966.39381.pdf). [Online]. Available:  
902 <https://doi.org/10.1001/jama.2023.8823>.

- 903 [55] D. Ayinde, N. Casartelli, and O. Schwartz, "Restricting hiv the samhd1 way: Through nu-  
904 cleotide starvation," *Nature Reviews Microbiology*, vol. 10, no. 10, pp. 675–680, Oct. 2012,  
905 ISSN: 1740-1534. DOI: 10.1038/nrmicro2862. [Online]. Available: <https://doi.org/10.1038/nrmicro2862>.
- 906  
907 [56] R. Wang *et al.*, "Sars-cov-2 restructures host chromatin architecture," *Nature Microbiology*,  
908 vol. 8, no. 4, pp. 679–694, Mar. 2023, ISSN: 2058-5276. DOI: 10.1038/s41564-023-01344-8.  
909 [Online]. Available: <https://doi.org/10.1038/s41564-023-01344-8>.
- 910 [57] N. Taefehshokr *et al.*, "Sars-cov-2 nsp5 antagonizes mhc ii expression by subverting histone  
911 deacetylase 2," *bioRxiv*, 2023. DOI: 10.1101/2023.02.10.528032. eprint: <https://www.biorxiv.org/content/early/2023/04/14/2023.02.10.528032.full.pdf>. [Online].  
912 Available: <https://www.biorxiv.org/content/early/2023/04/14/2023.02.10.528032>.
- 913  
914 [58] H. D. Desai, K. Sharma, J. V. Patoliya, E. Ahadov, and N. N. Patel, "A rare case of varicella-  
915 zoster virus reactivation following recovery from covid-19," *Cureus*, vol. 13, no. 1, e12423, Jan.  
916 2021, ISSN: 2168-8184. DOI: 10.7759/cureus.12423. [Online]. Available: <https://doi.org/10.7759/cureus.12423>.
- 917  
918 [59] T. Iwase *et al.*, "Mutation screening of the human clock gene in circadian rhythm sleep  
919 disorders," *Psychiatry Research*, vol. 109, no. 2, pp. 121–128, 2002, ISSN: 0165-1781. DOI:  
920 [https://doi.org/10.1016/S0165-1781\(02\)00006-9](https://doi.org/10.1016/S0165-1781(02)00006-9). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0165178102000069>.
- 921  
922 [60] R. Tedjasukmana, A. Budikayanti, W. R. Islamiyah, A. M. A. L. Witjaksono, and M. Hakim,  
923 "Sleep disturbance in post covid-19 conditions: Prevalence and quality of life," *Frontiers in  
924 Neurology*, vol. 13, 2023, ISSN: 1664-2295. DOI: 10.3389/fneur.2022.1095606. [Online].  
925 Available: <https://www.frontiersin.org/articles/10.3389/fneur.2022.1095606>.
- 926  
927 [61] C. Guissart *et al.*, "Dual molecular effects of dominant  $\text{jem}_\beta\text{rora}_\beta/\text{em}_\beta$  mutations cause two  
928 variants of syndromic intellectual disability with either autism or cerebellar ataxia," *The American  
929 Journal of Human Genetics*, vol. 102, no. 5, pp. 744–759, May 2018, ISSN: 0002-9297. DOI:  
930 [10.1016/j.ajhg.2018.02.021](https://doi.org/10.1016/j.ajhg.2018.02.021). [Online]. Available: <https://doi.org/10.1016/j.ajhg.2018.02.021>.
- 931  
932 [62] V. Venkataramani and F. Winkler, "Cognitive deficits in long covid-19," *New England Journal  
933 of Medicine*, vol. 387, no. 19, pp. 1813–1815, 2022. DOI: 10.1056/NEJMcibr2210069. eprint:  
934 <https://doi.org/10.1056/NEJMcibr2210069>. [Online]. Available: <https://doi.org/10.1056/NEJMcibr2210069>.
- 935  
936 [63] M.-D. Li *et al.*, " $\text{jem}_\beta\text{oi}/\text{em}_\beta$ -glcnac signaling entrains the circadian clock by inhibiting bmal1/clock  
937 ubiquitination," *Cell Metabolism*, vol. 17, no. 2, pp. 303–310, Feb. 2013, ISSN: 1550-4131. DOI:  
938 [10.1016/j.cmet.2012.12.015](https://doi.org/10.1016/j.cmet.2012.12.015). [Online]. Available: <https://doi.org/10.1016/j.cmet.2012.12.015>.

- 939 [64] J. Cavieres-Lepe and J. Ewer, "Reciprocal relationship between calcium signaling and circadian  
940 clocks: Implications for calcium homeostasis, clock function, and therapeutics," *Frontiers in*  
941 *Molecular Neuroscience*, vol. 14, 2021, ISSN: 1662-5099. DOI: 10.3389/fnmol.2021.666673.  
942 [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnmol.2021.666673>.
- 944 [65] M. Okubo, Y. Aoyama, and T. Murase, "A novel donor splice site mutation in the glycogen  
945 debranching enzyme gene is associated with glycogen storage disease type iii," *Biochemical and Biophysical Research Communications*, vol. 224, no. 2, pp. 493–499, 1996, ISSN: 0006-  
946 291X. DOI: <https://doi.org/10.1006/bbrc.1996.1055>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0006291X96910554>.
- 949 [66] N. Preisler *et al.*, "Exercise intolerance in glycogen storage disease type iii: Weakness or energy  
950 deficiency?" *Molecular Genetics and Metabolism*, vol. 109, no. 1, pp. 14–20, 2013, ISSN: 1096-  
951 7192. DOI: <https://doi.org/10.1016/j.ymgme.2013.02.008>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1096719213000607>.
- 953 [67] A. Leenhardt, I. Denjoy, and P. Guicheney, "Catecholaminergic polymorphic ventricular tachycardia," *Circulation: Arrhythmia and Electrophysiology*, vol. 5, no. 5, pp. 1044–1052, 2012. DOI:  
954 [10.1161/CIRCEP.111.962027](https://doi.org/10.1161/CIRCEP.111.962027). eprint: <https://www.ahajournals.org/doi/pdf/10.1161/CIRCEP.111.962027>. [Online]. Available: <https://www.ahajournals.org/doi/abs/10.1161/CIRCEP.111.962027>.
- 958 [68] C. Spuch *et al.*, "Efficacy and safety of lithium treatment in sars-cov-2 infected patients," *Frontiers in Pharmacology*, vol. 13, 2022, ISSN: 1663-9812. DOI: 10.3389/fphar.2022.850583.  
959 [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fphar.2022.850583>.
- 962 [69] E. C. Clarke, R. A. Nofchissey, C. Ye, and S. B. Bradfute, "The iminosugars celgosivir, castanospermine and UV-4 inhibit SARS-CoV-2 replication," *Glycobiology*, vol. 31, no. 4, pp. 378–  
963 384, Sep. 2020, ISSN: 1460-2423. DOI: 10.1093/glycob/cwaa091. eprint: <https://academic.oup.com/glycob/article-pdf/31/4/378/37761638/cwaa091.pdf>. [Online]. Available:  
964 <https://doi.org/10.1093/glycob/cwaa091>.
- 967 [70] A. Israel *et al.*, "Identification of drugs associated with reduced severity of covid-19: A case-control study in a large population," *medRxiv*, 2021. DOI: 10.1101/2020.10.13.20211953.  
968 eprint: <https://www.medrxiv.org/content/early/2021/03/24/2020.10.13.20211953.full.pdf>. [Online]. Available: <https://www.medrxiv.org/content/early/2021/03/24/2020.10.13.20211953>.
- 972 [71] A. Clemente-Moragón *et al.*, "Metoprolol in critically ill patients with covid-19," *Journal of the American College of Cardiology*, vol. 78, no. 10, pp. 1001–1011, 2021, ISSN: 0735-1097.  
973 DOI: <https://doi.org/10.1016/j.jacc.2021.07.003>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S073510972105590X>.

- 976 [72] H. Tomimoto, O. Takemoto, I. Akiguchi, and T. Yanagihara, "Immunoelectron microscopic  
977 study of c-fos, c-jun and heat shock protein after transient cerebral ischemia in gerbils,"  
978 *Acta Neuropathologica*, vol. 97, no. 1, pp. 22–30, Jan. 1999, ISSN: 1432-0533. DOI: 10.1007/  
979 s004010050951. [Online]. Available: <https://doi.org/10.1007/s004010050951>.
- 980 [73] P. Joshi *et al.*, "Host inducible-hsp70a1a is an irresistible drug target to combat sars-cov2  
981 infection and pathogenesis," *bioRxiv*, 2023. DOI: 10.1101/2023.05.05.539661. eprint: <https://www.biorxiv.org/content/early/2023/05/08/2023.05.05.539661.full.pdf>. [Online].  
982 Available: <https://www.biorxiv.org/content/early/2023/05/08/2023.05.05.539661>.
- 984 [74] R. W. Greene, T. E. Bjorness, and A. Suzuki, "The adenosine-mediated, neuronal-glial, home-  
985 static sleep response," *Current Opinion in Neurobiology*, vol. 44, pp. 236–242, 2017, Neu-  
986 robiology of Sleep, ISSN: 0959-4388. DOI: <https://doi.org/10.1016/j.conb.2017.05.015>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959438817300466>.
- 989 [75] P. Sassone-Corsi, "The cyclic amp pathway," *Cold Spring Harbor Perspectives in Biology*,  
990 vol. 4, no. 12, 2012. DOI: 10.1101/cshperspect.a011148. eprint: <http://cshperspectives.cshlp.org/content/4/12/a011148.full.pdf+html>. [Online]. Available: <http://cshperspectives.cshlp.org/content/4/12/a011148.short>.
- 993 [76] P. ENDRES and T. ROEREN, "Lack of a Diurnal Plasma Adenosine 3,5-Monophosphate  
994 Rhythm," *The Journal of Clinical Endocrinology Metabolism*, vol. 48, no. 5, pp. 872–873,  
995 May 1979, ISSN: 0021-972X. DOI: 10.1210/jcem-48-5-872. eprint: <https://academic.oup.com/jcem/article-pdf/48/5/872/10513832/jcem0872.pdf>. [Online]. Available:  
996 <https://doi.org/10.1210/jcem-48-5-872>.
- 998 [77] Y. Sassi *et al.*, "Cardiac myocyte-secreted camp exerts paracrine action via adenosine receptor  
999 activation," *The Journal of Clinical Investigation*, vol. 124, no. 12, pp. 5385–5397, Dec. 2014.  
1000 DOI: 10.1172/JCI74349. [Online]. Available: <https://doi.org/10.1172/JCI74349>.
- 1001 [78] J. Leroy *et al.*, "Spatiotemporal dynamics of beta-adrenergic camp signals and l-type cajup{2+i}/sup{i}  
1002 channel regulation in adult rat ventricular myocytes," *Circulation Research*, vol. 102, no. 9,  
1003 pp. 1091–1100, 2008. DOI: 10.1161/CIRCRESAHA.107.167817. eprint: <https://www.ahajournals.org/doi/pdf/10.1161/CIRCRESAHA.107.167817>. [Online]. Available: <https://www.ahajournals.org/doi/abs/10.1161/CIRCRESAHA.107.167817>.
- 1006 [79] B. Tenner *et al.*, "Spatially compartmentalized phase regulation of a  $\text{Ca}^{2+}$ -camp-pka oscillatory  
1007 circuit," *eLife*, vol. 9, S. Mayor, J. A. Cooper, and A. Nadler, Eds., e55013, Nov. 2020, ISSN:  
1008 2050-084X. DOI: 10.7554/eLife.55013. [Online]. Available: <https://doi.org/10.7554/eLife.55013>.
- 1010 [80] N. Grotehans *et al.*, "Ribonucleotide synthesis by nme6 fuels mitochondrial gene expression,"  
1011 *The EMBO Journal*, vol. n/a, no. n/a, e113256, DOI: <https://doi.org/10.15252/embj.2022113256>. eprint: <https://www.embopress.org/doi/pdf/10.15252/embj.2022113256>.  
1012 [Online]. Available: <https://www.embopress.org/doi/abs/10.15252/embj.2022113256>.

- 1014 [81] E. L. Hatcher *et al.*, “Virus Variation Resource – improved response to emergent viral out-  
1015 breaks,” *Nucleic Acids Research*, vol. 45, no. D1, pp. D482–D490, Nov. 2016, ISSN: 0305-1048.  
1016 DOI: 10.1093/nar/gkw1065. eprint: <https://academic.oup.com/nar/article-pdf/45/D1/D482/8846723/gkw1065.pdf>. [Online]. Available: <https://doi.org/10.1093/nar/gkw1065>.
- 1017  
1018 [82] W. C. Forsythe, E. J. Rykiel, R. S. Stahl, H.-i. Wu, and R. M. Schoolfield, “A model com-  
1019 parison for daylength as a function of latitude and day of year,” *Ecological Modelling*, vol. 80,  
1020 no. 1, pp. 87–95, 1995, ISSN: 0304-3800. DOI: [https://doi.org/10.1016/0304-3800\(94\)00034-F](https://doi.org/10.1016/0304-3800(94)00034-F). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/030438009400034F>.
- 1021  
1022  
1023 [83] E. Guidotti and D. Ardia, “Covid-19 data hub,” *Journal of Open Source Software*, vol. 5,  
1024 no. 51, p. 2376, 2020. DOI: 10.21105/joss.02376.
- 1025  
1026 [84] E. Guidotti, “A worldwide epidemiological database for covid-19 at fine-grained spatial reso-  
1027 lution,” *Scientific Data*, vol. 9, no. 1, p. 112, 2022. DOI: 10.1038/s41597-022-01245-1.
- 1028  
1029 [85] I. Berry *et al.*, “A sub-national real-time epidemiological and vaccination database for the  
1030 covid-19 pandemic in canada,” *Scientific Data*, vol. 8, no. 1, Jul. 2021. DOI: 10.1038/s41597-  
1031 021-00955-2. [Online]. Available: <https://doi.org/10.1038/s41597-021-00955-2>.
- 1032  
1033 [86] E. Dong, H. Du, and L. Gardner, “An interactive web-based dashboard to track covid-19 in  
1034 real time,” *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, May 2020, ISSN: 1473-  
1035 3099. DOI: 10.1016/S1473-3099(20)30120-1. [Online]. Available: [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1).
- 1036  
1037 [87] S. de Salud Mexico, “Información referente a casos covid-19 en méxico,” May 2022. [Online].  
1038 Available: <https://datos.gob.mx/busca/dataset/informacion-referente-a-casos->  
1039 [covid-19-en-mexico](https://covid-19-en-mexico).
- 1040  
1041 [88] M. de Salud Argentina, “Covid-19. casos registrados en la república argentina,” May 2022.  
1042 [Online]. Available: <http://datos.salud.gob.ar/dataset/covid-19-casos-registrados->  
1043 [en-la-republica-argentina](https://en-la-republica-argentina).
- 1044  
1045 [89] M. de Ciencia and M. de Salud, “Dp3 - casos totales por región,” May 2022. [Online]. Available:  
1046 <https://github.com/MinCiencia/Datos-COVID19/tree/master/output/producto3>.
- 1047  
1048 [90] M. de Salud - MINSA, “Casos positivos por covid-19,” May 2022. [Online]. Available: <https://www.datosabiertos.gob.pe/dataset/casos-positivos-por-covid-19-ministerio-de-salud-minsa>.
- 1049  
1050 [91] P. de Datos Abiertos Gobierno de Colombia, “Casos positivos de covid-19 en colombia,” May  
1051 2022. [Online]. Available: <https://www.datos.gov.co/Salud-y-Proteccion-Social/Casos-positivos-de-COVID-19-en-Colombia/gt2j-8ykr/data>.
- 1052  
1053 [92] S. W. D. Center, “The international sunspot number,” *International Sunspot Number Monthly  
1054 Bulletin and online catalogue*, [Online]. Available: <http://www.sidc.be/silso/>.

- 1050 [93] E. Richard, “Tsis sim level 3 solar spectral irradiance 24-hour means v09,” *Goddard Earth*  
1051 *Sciences Data and Information Services Center (GES DISC)*, 2022. doi: DOI:10.5067/TSIS/  
1052 *SIM/DATA318*. [Online]. Available: [https://disc.gsfc.nasa.gov/datacollection/TSIS\\_SSI\\_L3\\_24HR\\_09.html](https://disc.gsfc.nasa.gov/datacollection/TSIS_SSI_L3_24HR_09.html).
- 1053  
1054 [94] G. E. S. Data and I. S. C. ( DISC), “Aqua/airs l3 daily standard physical retrieval (airs-only)  
1055 1 degree x 1 degree v7.0,” *AIRS project (2019)*, 2023. [Online]. Available: <https://doi.org/10.5067/U03Q64CTTS1U,.>
- 1056