 doi:10.20944/preprints202210.0217.v1

Article

Nipah Pandemic Preparedness: Machine Leaning-Derived Whole Brain Organoid Simulations for Identifying Optimal Vaccine Candidates against the Nipah Virus

Sally Ezra and Wayne R Danter

123Genetix, 11280 Riverside Drive E, Windsor, Canada, N8P 1A4

* Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials)

ABSTRACT: The evolving global SARS-C0V-2 pandemic emphasizes how unprepared we are for the emergence of the next lethal viral pathogen. The World Health Organization has created a list of such potential candidates and identified Nipah virus infection as the highly lethal prototypic member of that list. Building on our previous viral pandemic preparedness research into SARS-CoV-2 we have created computer simulations of Nipah virus infection complicated by encephalitis, the most common cause of Nipah-associated mortality. In the current experiments, we first created updated simulations of wild-type whole-brain organoids (aiWBO). Once validated the aiWBO was infected with the simulated Nipah virus genome. The Nipah encephalitis simulations (aiWBO-NiV) were then used to identify the optimal single, double, and triple protein combinations for candidates as potential targets for vaccine development. Our data suggest that the use of multi-viral proteins/epitopes chimera is the most promising Nipah vaccine approach and that employing artificial intelligence-guided antigen design can be used to rapidly identify vaccine candidates for pandemic preparedness and rapid response.

Keywords: Brain Organoids; Artificial Intelligence; Machine Learning; infectious diseases; Nipah

Introduction

SARS-CoV-2 is the "canary in the coal mine" and we continue to ignore the warning at our peril. The severe consequences of the global SARS-CoV-2 pandemic were due to a lack of pandemic preparedness [1]. Thus, it is crucial that we begin to develop strategies to prepare the world, for when the next deadly viruses emerge. Recently, an outbreak of Nipah virus (NiV) in India has raised the question of whether we should start to consider it as a future threat and look to develop NIPAH pandemic mitigating measures including a NiV vaccine [2].

According to the World Health Organization (WHO, viral diseases like Avian flu, Swine flu, Middle East respiratory syndrome coronavirus (MERS-CoV), Severe acute respiratory syndrome (SARS), Crimean-Congo haemorrhagic fever (CCHF), Ebola virus disease and Marburg virus disease, Lassa fever, Nipah virus (NiV), Henipaviral diseases, Rift Valley fever, Zika, and Disease X, a new, previously unknown virus; pose considerable risk of an international public health emergency, when these spread rapidly [3]. In the past 5 years emergency situations have been created by SARS-CoV-2, Ebola and Zika viruses [4]. COVID-19, the disease caused by SARS-CoV-2 infection, emerged in 2020 and produced a global mobilization of researchers seeking innovative technologies for developing efficient and rapid diagnostics, vaccines, drug targets and therapies for combating COVID-19. These efforts continue to save the lives of millions of people across the globe. Unfortunately, sporadic Nipah virus outbreaks continue to increase concern in the research community and the public.

Nipah virus infection in humans causes a range of clinical presentations, from asymptomatic infection (subclinical) to acute respiratory infection and fatal encephalitis

 doi:10.20944/preprints202210.0217.v1

[3, 4]. The high fatality rate associated with Nipah disease and the lack of any effective treatments and vaccines, constitute an impending global threat [5]. The disease was first recognized in Asia in 1999 and is associated with a death rate of up to 70% in those infected [6]. Fortunately, a validated Nipah viral genome is available to researchers [7]. Given that the WHO has identified Nipah virus as a prototype for a future pandemic, there is a pressing need to develop effective Nipah vaccines to prevent a widespread outbreak [3]. Towards this goal, the present study evaluates an updated version of our machine-learning platform DeepNEU (v6.6) for creating computer simulations of artificially induced whole brain organoid (aiWBO) derived from artificially induced human pluripotent stem cells (aiPSCs) for helping to prevent a global Nipah virus epidemic and pandemic. The uninfected, wild type aiWBO was then exposed to infection with simulated Nipah virus (NiV) to specifically simulate Nipah virus encephalitis, the most common cause of death associated with Nipah virus infection. Finally, the aiWBO-NiV simulations were applied to NiV vaccine discovery and optimization. The genomic and phenotypic profiles of wild-type aiWBO (uninfected) and aiWBO-NiV encephalitis simulations were validated using the currently available experimental wet lab and Nipah genomic data (a single stranded RNA molecule with linear topology containing 18246 nucleotide base pairs). For example, we used the Nipah genome data (GenBank accession number: NC_002728.1) which was retrieved from NCBI Genome database (https://www.ncbi.nlm.nih.gov/genome/) to simulate Nipah-mediated encephalitis.

Methods

Our DeepNEU (v6.6, 123Genetix Inc., London and Windsor, Canada) stem cell and organoid simulation platform is a literature-validated unsupervised deep-machine learning system with features of fully connected recurrent neural networks (RNN), cognitive maps (CM), support vector machines (SVM) and evolutionary systems (GA). The essential component of the DeepNEU platform is a large and growing database of relationships between concepts that are derived from gene and protein cell signaling pathways data assembled from the peer reviewed literature. The information in the database is stored in the form of a square matrix and matrix with values that range from -1 for a maximally negative relationship to +1 for a maximally positive relationship. In this system 0 is used to represent an arbitrary baseline value. This fuzzy logic (FL) representation can be converted to a neutrosophic logic (NL) by representing truly unknown relationships with the letter I for indeterminant. An initial state input vector is used to represent the state of the system at time zero. Iterative sparse vector matrix multiplication cycles continue until a new system-wide, steady state is achieved. The initial state vector can be modified to represent a wide range of genetic disorders by turning on or off any number or combination of gene, protein or phenotypic concepts. Our machine learning platform uses unsupervised learning combined with a large amount of data and an early stopping form of regularization to minimize overfitting the data. In this case, by unsupervised learning we mean that there are no data profiles associated with specific outcomes and the system instead learns relationships between individual concepts and not specific outcomes. In our approach, unsupervised learning would be more akin to clustering and regression than classification. The DeepNEU database generation and validation of artificially induced whole-brain organoids (aiWBO) have been described in detail in earlier peer-reviewed publications [8, 9].

In this project, the upgraded DeepNEU database (v6.6) was generated from an earlier database (v6.2) by adding 774 new genotypic and phenotypic features and 5000 new relationships that were relevant for developing whole-brain organoids and NiV virus encephalitis. These new features produced improved simulation of specific brain organoid cell types, brain regions, the blood–brain barrier (BBB) and microcirculation. Briefly, the previous DeepNEU database (v6.2) contained 4660 proteins or genes and 43,498 nonzero relationships, while the current database (v6.6) contains 5434 proteins or genes and 48,487 nonzero relationships, gene/proteins, or phenotypic inputs and outputs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

The Whole-Brain Organoid Simulations, Using NEUBOrg (DeepNEU v6.6)

The primary goal of this project was to extend our recently developed NEUBOrg platform and validated aiWBO simulations using the current version of the DeepNEU (v6.6) machine learning platform. As reported previously [8, 9], we used a simulated approach like the one published by Yamanaka's group in 2007 [10] to generate induced pluripotent stem cells (iPSCs) from human fibroblasts. Several changes to the original 2007 protocol were required to promote differentiation from artificially induced pluripotent stem cells (aiPSC) to aiWBO. As before, the modified simulation protocol began with the activation of the four Yamanaka transcription factors, namely octamer-binding transcription factor 4 (OCT4), cellular myelocytomatosis oncogene (cMYC), kruppel-like factor 4 (KLF4) and sex- determining region Y-box 2 (SOX2). Instead of the iPSC media used by Yamanaka [10], we substituted a simulated B27 neural media with supplementary zinc, ascorbic acid, and Vitamin D3 in the presence of doxycycline, normal levels of ambient oxygen (21%) and increased ambient CO2 (5%). The original Yamanaka protocol was carried out under hypoxic conditions and normal levels of CO2. As detailed previously, the reprogramming of aiPSC into aiWBO was conducted in a simulated rotating bioreactor [8]. The aiWBO simulations were carried out in triplicate and the simulation protocol employed is summarized in Table 1.

Table 1. Unguided protocol for simulating the transformation of aiPSC to artificially induced whole-brain organoids (aiWBO).

Simulation Summary	Components (Simulated)
Yamanaka (2007) transcription factors	OCT4, cMYC, KLF4 and SOX2 turned ON
B27 neural media	biotin, amino acids, ascorbate, catalase, cortisol, basic fibroblast growth factor (FGF2/bFGF), glutathione, albumin, insulin, SOD1(Cu/Zn), MnSOD/SOD2, progesterone, retinol/Vitamin A, thyroid hormones (T3/T4), transferrin, Vitamin E/Tocopherol, L-carnitine locked ON
Supplements	Zinc, Vitamin D3 and doxycycline turned ON
Rotating bioreactor (optimized)	B27 media + [CO2] = 5%, [O2] = 21%, glucose, temperature = 37 degrees C locked ON and high shear forces locked OFF

The final aiWBO simulations reproduced several relevant brain cell types (N=9), rostral-caudal brain regions, ventral dorsal regions, six horizontal layers of the cerebral cortex and the four layers of the cerebellar cortex. As before, the spinal cord was not simulated in these experiments. The main cell types included neural precursor cells (Radial Glial/NSC) and neurons, astrocytes, oligodendrocyte precursor cells (OPC), oligodendrocytes, interneurons, and microglial cells. In addition, the system was also able to simulate endothelial cells and pericytes, which are important components of the BBB. The forebrain, midbrain and hindbrain were simulated as part of the rostral caudal regions, and the ventral dorsal regions and the ventral forebrain as well as six layers of the cerebral cortex were also successfully simulated. For more details on the genotypic and phenotypic features that we have used to build the NEUBOrg platform to simulate aiWBO please refer to [11-27].

The NEUBOrg, aiWBO Simulations Applied to NiV Virus Infection

The validated and unguided NEUBOrg platform was also used to simulate a whole-brain organoid infected by the NiV virus. Briefly, the NEUBOrg was upgraded by including previously reported NiV virus specific phenotypic and genotypic features [28-32]. The simulation process was otherwise identical to that outlined above for the development of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

the wild type aiWBO simulations (aiWBO–WT). As before, all aiWBO–NiV and aiWBO–WT simulation experiments were carried out in triplicate.

In these experiments, we exposed the aiWBO to the simulated NiV virus. For this simulated infection, the concept of initial viremia was activated (turned on). The viremia activates the viral life cycle beginning with the interaction of the F and G viral proteins with the cellular receptor protein Eph2 and ending with exocytosis of new viral particles which completes the life cycle by contributing new viral particles to the ongoing viremia [28-32]. The NiV genome consists of six structural genes that produce at least nine proteins. As described previously, the 9 gene/protein expression profile was compared with the uninfected aiWBO simulations to assess the validity of simulated NiV infection. All genes and proteins studied were expressed in the aiWBO-NiV, but not aiWBO simulations. The probability that all (N = 9) of these aiWBO-NiV simulation outcomes were correctly predicted by chance alone using the binomial test is 0.0241.

To evaluate the performance of the aiWBO–NiV profile compared to the aiWBO– WT profile, we used several phenotypic features (N=11) that were reported to reflect the NiV infection profile [2, 28-34]. Features with negative inputs were used to simulate inhibition of the feature, while features with positive inputs were used to promote the feature. A compilation of the substantial number and status of inputs that constituted the aiWBO–NiV feature profile is presented in Table 2.

Table 2. NiV virus infection (aiWBO-NiV) phenotypic feature profile.

NiV Features (N = 11)	Genotypic/Phenotypic Feature Inputs
Astrocyte activation	N = 32 (6 negatives and 26 positives)
Microglial activation	N = 83 (24 negatives and 59 positives)
IgM	N = 5 (0 negatives and 5 positives)
Endothelial infection	N = 6 (0 negatives and 6 positives)
Neuroinflammation	N = 58 (15 negatives and 43 positives)
Seizures	N = 21 (5 negatives and 9 positives)
Vasculitis	N = 4 (0 negative and 4 positives)
Viral Load	N = 4 (1 negative and 3 positives)
IFNa/b	N = 44 (17 negative and 27 positives)
ISGF3	N = 6 (3 negative and 3 positives)
ISGs	N = 14 (3 negative and 11 positives)

NEUBOrg Platform Statistical Analysis

Consistent with our previous research [7], we used the unbiased binomial test to conduct statistical analysis between the data that were predicted by aiWBO–WT or aiWBO–NiV simulations vs peer-reviewed published data. This statistical analysis was selected as it can compensate for prediction bias. To calculate the significance of difference between groups (e.g., aiWBO–WT vs aiWBO–NiV), we used the two-tailed unpaired t-test with unequal variances (Welch's t-test).

Discovering Potential Vaccine Candidates for Mitigating NiV virus Encephalitis

Once the aiWBO-NiV simulations were validated using available published wet lab data, that the system had not seen before, the protein targets (N=9) derived from the NiV genome (proteins N, P, W, V, C, M, F, G and L) were evaluated as potential immunogenic components of monovalent and/or polyvalent vaccine candidates [2, 28-34]. To simulate the untreated disease state, we included a placebo. The placebo was created using a treatment vector composed of all zeros. In other words, no concepts were turned on.

The impact of potential candidates was assessed using the previously described NiV disease profile which was equivalent to the placebo or untreated profile. As outlined above the phenotypic disease (encephalitis) profile (N = 11 factors) included Astrocytes-Activated, Endothelial cell Infection, IFNa/b, IgM, ISGF3, ISGs/ISG15/ISG3, Microglia-Activated, Neuroinflammation, NiV-Vasculitis, NiV-Viral load and Seizures.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

A validated distance measure, the angular cosine distance (ACD), was used to evaluate all treatment effects relative to placebo. The ACD is a statistical metric for representing the distance between two or more real valued vectors. The ACD is based on a widely used similarity measure, the cosine similarity (CS). Once calculated, the CS can be converted to ACD using the formula ACD = arccosine (CS)/Pi. As distance from the reference vector increases the ACD increases towards 1, while decreasing distance from the reference moves the ACD toward 0. The ranking and final selection of candidates used a two-step process. First, we applied the ACD distance metric to rank the distance of treatments relative to the aiWBO-WT from smaller to larger values. Second, statistically significant differences between aiWBO-NiV/placebo and treatment effects were ranked using the non-parametric Wilcoxon test and a p value < 0.05 was interpreted to indicate that observed differences were unlikely to be due to chance alone.

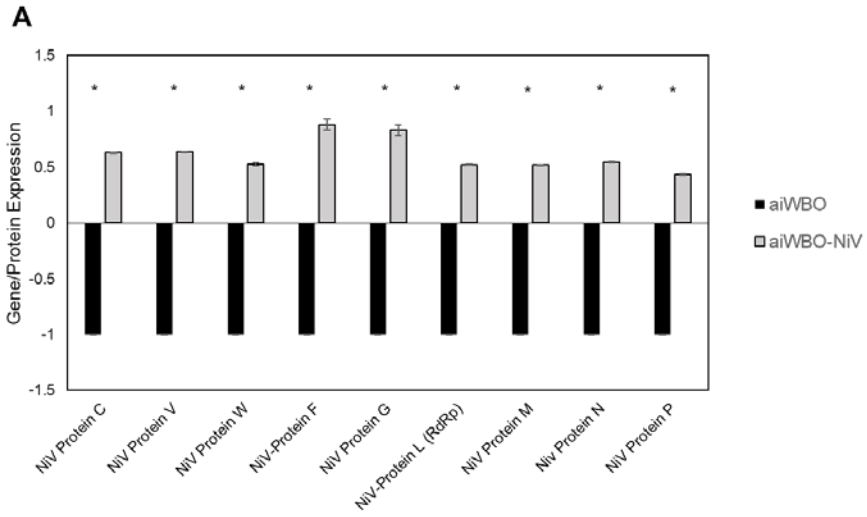
Results

The NEUBOrg (v6.6) Platform Specification

As described in the method section, the updated NEUBOrg/DeepNEU database contained a large amount of phenotypic and genotypic data, which generated a large amount of information flowing through each node in the CM/RNN. On average, each node in the network initially had 9.0 inputs and 9.0 outputs. The pre-simulation probability of a positive outcome prediction was 0.661 and the pre-test probability of a negative prediction was therefore 0.339. To minimize any prediction biases, this system bias was used when applying the binomial test to all simulation outcomes. In the current version of the platform, the spinal cord was not simulated.

1. Disease Modeling Simulations of the aiWBO–NiV virus encephalitis

The expression of Nipah virus gene/protein profile in Nipah infected whole-brain organoid, aiWBO-NiV relative to uninfected aiWBO-WT were accurately predicted as presented in (Figure 1A). Additionally, the expression of the 11 features that represent the NiV disease profile relative to uninfected aiWBO-WT were also accuarately predicted (Figure 1B). The probability that the expression of all NiV gene/protein concepts were predicted by chance alone is 0.011 (binomial test). Statistical analysis indicated that there were significant changes in all components of the simulated aiWBO–NiV disease profile with reference to the aiWBO–WT profile. For example, Astrocytes-Activated, Endothelial cell Infection, IgM, Microglia-Activated, Neuroinflammation, NiV Vasculitis, NiV-Viral load and Seizures were all increased ($p<0.05$, Welch's t-test) while indicators of IFNa/b signaling including IFNa/b, ISGF3, and ISGs/ISG15/ISG3 were all decreased ($p<0.05$).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

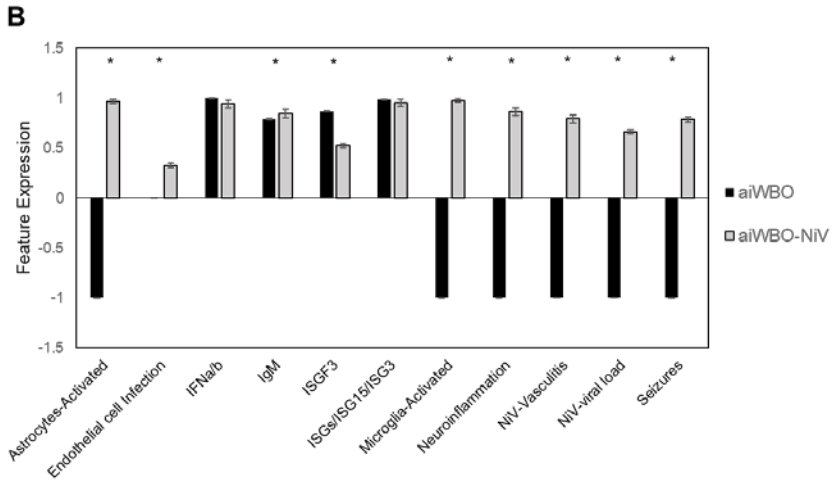


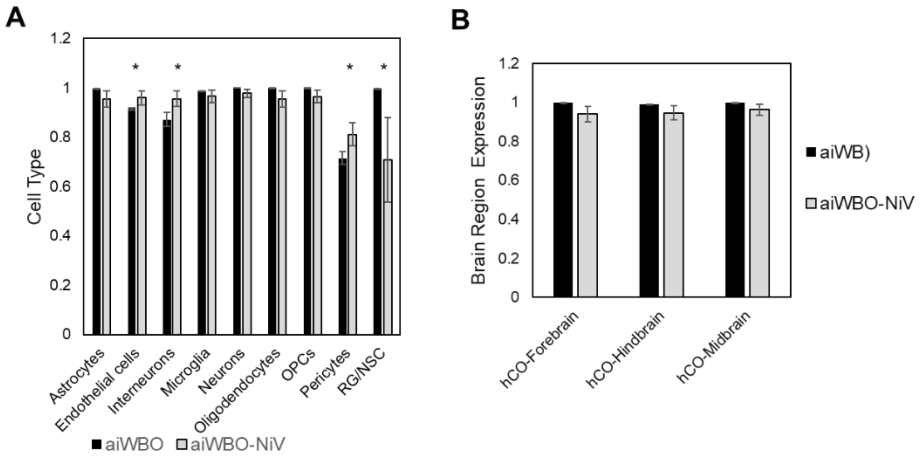
Figure 1. Simulation of Nipah infection and Nipah-mediated Encephalitis profile, A, Gene/protein expression profile in Nipah infected aiWBO-NiV reference to uninfected aiWBO-WT. B, comparative simulation of Encephalitis profile feature in aiWBO–WT vs aiWBO–NiV. The vertical y-axes represent the semiquantitative levels of concepts that are estimated by NEUBOrg relative to an arbitrary baseline where 0 = baseline, 1 = maximum expression or presence and -1 = minimal expression level or presence. The horizontal x-axes represent the individual aiWBO concepts being simulated. Data represent means of three experiments ± 95% confidence interval. All p values from the Welch's t test. *p<0.01.

2. The aiWBO-NiV and aiWBO Wild Type Simulations

Consistent with previous results [8, 9], the aiWBO-NiV and aiWBO-WT simulations converged quickly (after 23 and 32 iterations respectively) to new system-wide steady states without evidence of overtraining after 1000 iterations. The simulation results of aiWBO–NiV and aiWBO–WT are presented below:

2.1. Simulation of Neural Cell Types in aiWBO–NiV vs aiWBO-WT

The aiWBO–NIV and aiWBO–WT simulations successfully predicted the expression of all nine common neural cell types that are found in human brain organoids. The probability that all (N = 9) cell types were predicted by chance alone is 0.023 (binomial test). A statistical comparison using the Welch test also indicated that there were significant increases in the expression of endothelial cells, interneurons and pericytes, while astrocytes, neurons, oligodendrocytes, OPC and RG/NSC were significantly decreased in the aiWBO–NiV simulations (all p values < 0.05). No significant difference was detected in resting microglial cells. These results are summarized in Figure 2A.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

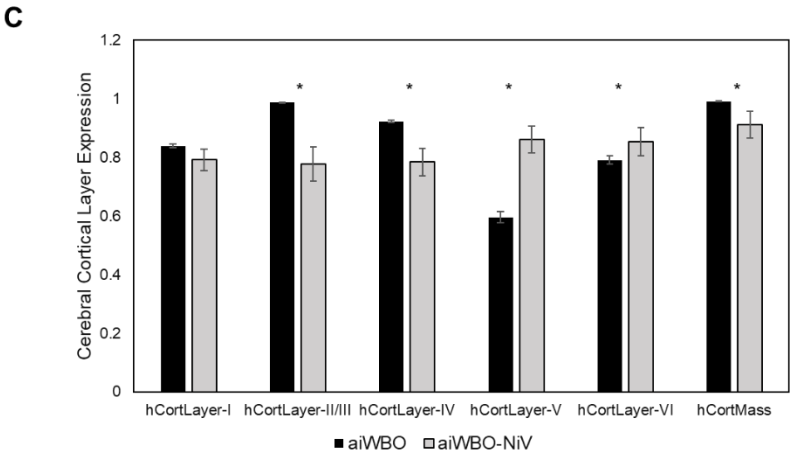


Figure 2. Simulation of cell types, brain regions and layers in aiWBO vs aiWBO–NiV. A, Comparison of wild type artificially induced whole-brain organoid (aiWBO/WT) vs Nipah Encephalitis aiWBO (NiV) cell types. B, Comparison of primitive brain region concepts in WT vs NiV simulations, C, Comparison of cerebral cortical layers in WT vs NiV simulations. The vertical y-axes represent the semiquantitative levels of concepts that are estimated by NEUBOrg relative to an arbitrary baseline where 0 = baseline, 1 = maximum expression or presence and -1 = minimal expression level or presence. The horizontal x-axes represent the individual aiWBO concepts being simulated. Data represent means of three experiments ± 95% confidence interval. All p-values from the Welch's t test. *p<0.01.

2.2. Simulation of Rostral–Caudal Brain Regions in aiWBO–NIV vs aiWBO-WT

The aiWBO–NiV and aiWBO–WT simulations successfully predicted the expression of all three rostral–caudal brain regions of human brain organoids, namely, the forebrain, midbrain and hindbrain. The probability that the expression of all (N = 3) regions were predicted by chance alone is 0.285 (binomial test). There was a significant decrease in the expression of the forebrain, hindbrain and midbrain in the aiWBO–NiV simulations with reference to aiWBO–WT ($p<0.05$, Welch test). These results are summarized in Figure 2B.

2.3. Simulation of Cerebral Cortical Layers in aiWBO–NiV vs aiWBO-WT

The aiWBO–NiV and aiWBO-WT simulations successfully predicted the expression of all six cerebral cortical layers found in human brain organoids ($p = 0.083$, binomial test). A statistical analysis using the Welch t-test indicated that there was a significant decrease ($p < 0.05$) in the expression of layers 1,2/3, and 4 with increases observed ($p < 0.05$) in layers 5 and 6. These changes combined to produce a significant decrease ($p<0.01$) in estimated cerebral cortical mass in the NiV simulations. These results are summarized in Figure 2C.

2.4. Simulation of Cerebellar Cortical Layers in aiWBO–NiV vs aiWBO-WT

The aiWBO–NiV and aiWBO–WT simulations successfully predicted the expression of all four cerebellar cortical layers found in human brain organoids. ($p = 0.188$, binomial test). No significant changes in the expression of two of the four cerebellar cortical layers was observed. There was a significant increase noted in the Molecular Layer in the aiWBO–NiV relative to the aiWBO–WT simulations ($p<0.001$, Welch test) and a decrease in the Granular Layer ($p<0.05$). These results are summarized in Figure 3A.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

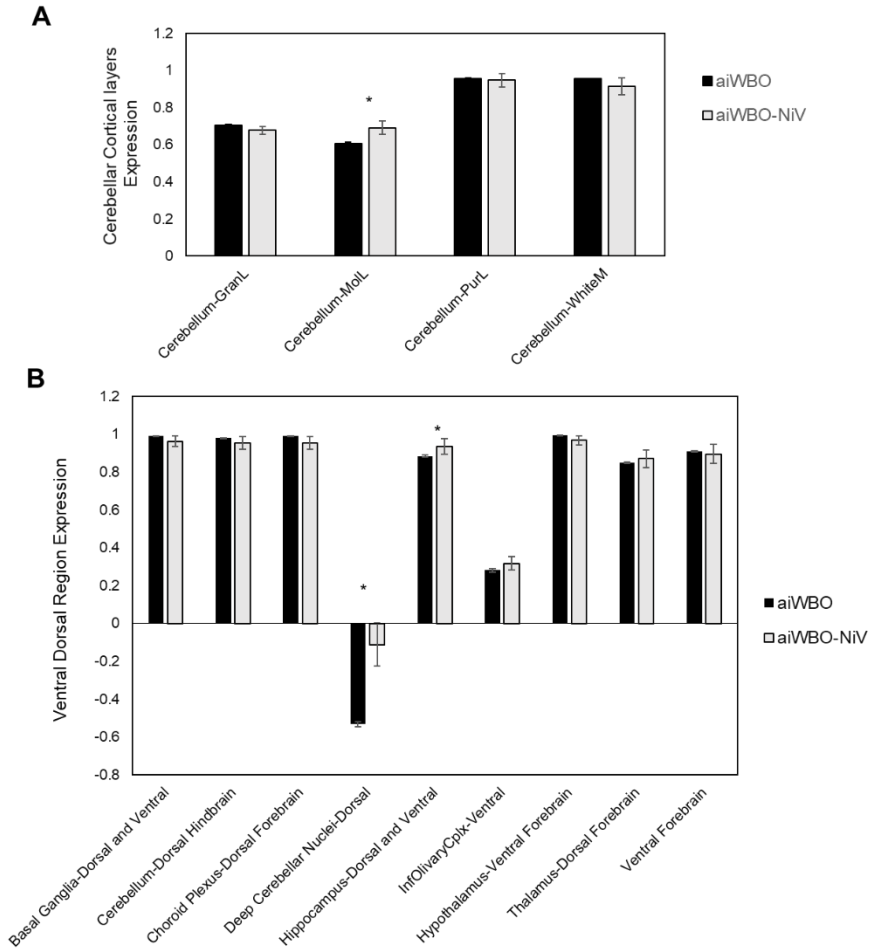


Figure 3. Expression of Cerebellar cortical Layers and Ventral Dorsal Regions in Wild Type and Nipah (NiV)-Encephalitis brain organoid Simulation. A, Comparison of cerebellar cortical layers expression in aiWBO and aiWBO-NiV simulations. B, Comparison of ventral Dorsal Region expression in WT vs NiV simulations. The vertical y-axes represent the semiquantitative levels of concepts that are estimated by NEUBOrg relative to an arbitrary baseline where 0 = baseline, 1 = maximum expression or presence and -1 = minimal expression level or presence. The horizontal x-axes represent the individual aiWBO concepts being simulated. Data represent means of three experiments ± 95% confidence interval. All p values from the Welch's t test. *p<0.01.

2.5. Simulation of Ventral–Dorsal Brain Regions in aiWBO–WT vs aiWBO-NiV

The aiWBO–WT and aiWBO–NiV simulations successfully predicted the expression of all nine ventral (anterior)–dorsal (posterior) representative regions that are commonly found in wet lab human brain organoids. The probability that the expression of all (N = 9) regions were predicted by chance alone is 0.023 (binomial). There were no significant differences ($p > 0.05$) in the expression of ventral forebrain regions in the aiWBO–WT vs the aiWBO–NiV simulations (Welch's t test). Several other regions were increased in the aiWBO–NiV vs the aiWBO–WT simulations ($p < 0.01$). The increased regions include the basal ganglia (dorsal and ventral), cerebellum (dorsal), choroid plexus (dorsal), hypothalamus (ventral), inferior olivary complex (ventral) and thalamus (dorsal). The decreased regions in the aiWBO– NiV vs the aiWBO–WT simulations ($p < 0.01$) include the deep cerebellar nuclei (dorsal) and the hippocampus (dorsal and ventral). These results are summarized in Figure 3B.

2.6. Simulation of Microcirculation in aiWBO–NiV vs aiWBO-WT

The aiWBO–NiV and aiWBO–WT simulations correctly predicted the expression of all seven elements of a microcirculation found in human brain organoids ($p = 0.055$, binomial test). Observed increases in the intracellular O2 concentration and cerebral spinal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

fluid in the aiWBO–NiV simulations suggest a degree of microcirculation improvement ($p<0.01$, Welch's t-test). While there was an increase in endothelial cells ($p<0.01$) this increase was associated with marked increase in endothelial dysfunction ($p<0.0001$). We also observed a decrease in arterial and venous endothelial cells ($p<0.05$) consistent with marked endothelial dysfunction ($p<0.001$). There were no significant changes in Lymphatic endothelial cells in the aiWBO–NIV relative to the aiWBO–WT simulations ($p=0.77$). These results are summarized in Figure 4A.

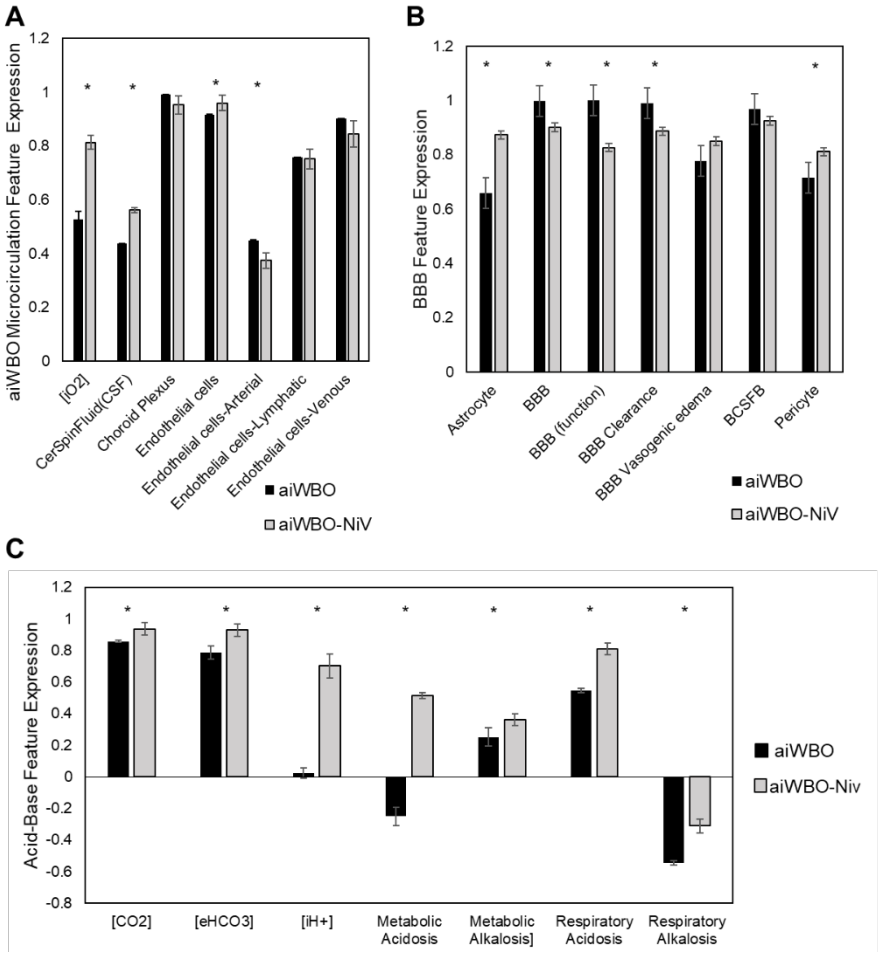


Figure 4. Comparative analysis of Microcirculation, Blood-Brain Barrier and Acid-Base feature in Nipah (NiV) encephalitis vs WT. A, comparative simulation analysis of microcirculation features in aiWBO–WT vs aiWBO–NiV. B, Comparative simulations analysis of Blood-Brain Barrier (BBB) in aiWBO vs aiWO-NiV. C, comparative simulation analysis of Acid-Base features in aiWBO–WT vs aiWBO–NiV. The vertical y-axes represent the semiquantitative levels of concepts that are estimated by NEUBOrg relative to an arbitrary baseline where 0 = baseline, 1 = maximum expression or presence and -1 = minimal expression level or presence. The horizontal x-axes represent the individual aiWBO concepts being simulated. Data represent means of three experiments ± 95% confidence interval. All p values from the Welch's t test. *$p<0.01$.

2.7. Simulation of Blood–Brain Barrier in aiWBO–NiV vs aiWBO-WT

The aiWBO–NiV and aiWBO–WT successfully predicted the expression of all seven features that represent the BBB profile in human brain organoids ($p = 0.055$, binomial test). Statistical analysis indicated that there were significant increases in the astrocyte and pericyte components of the BBB ($p<0.01$, Welch's t-test). The expression of the BBB itself was decreased ($p < 0.01$) in association with decreased BBB clearance ($p < 0.001$) and a significant increase in vasogenic brain edema ($p<0.01$) was observed, all consistent with significant BBB dysfunction ($p<0.001$). A non-significant decrease in the blood CSF barrier (BCSFB) was observed ($p<0.063$). These results are summarized in Figure 4B.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

2.8. Simulation of Acid-Base Status in aiWBO–NiV vs aiWBO-WT

The expression of all seven features that represent compensated metabolic alkalosis and respiratory acidosis in human brain organoids were predicted successfully and the data are summarized in Figure 4C.

The probability that the expression of all (N = 7) concepts associated with the existence of a complex respiratory and metabolic acidosis were predicted by chance alone is 0.055 (binomial test). There were significant changes ($p < 0.01$, Welch's t-test) in all components of the aiWBO–NiV simulations with reference to aiWBO–WT. The presence of a severe acidosis is confirmed by a marked increase in intracellular [H+] concentration in aiWBO–NiV ($p<0.001$). In summary, these data suggest a poorly compensated complex respiratory and metabolic acidosis in the aiWBO–NiV simulations.

2.9. Cellular Stress and Neuronal Cell Death in aiWBO-NiV vs aiWBO-WT simulations

The aiWBO–WT and aiWBO–NiV simulations predicted the expression of six markers of cellular stress and neuronal cell death that are commonly observed in in vitro human brain organoids. A statistical analysis using the Welch's t-test indicated that there were significant changes in all but one (i.e., PGK1/PK-1) of these markers. ER Stress, ARCN1, GORASP2 and Neuron cell death were all increased ($p<0.001$, Welch's t-test) while Glycolysis was decreased ($p<0.05$). These results are summarized in Figure 5. Overall, there is a significant increase in both observed cellular stress and neuronal cell death in the encephalitis simulations.

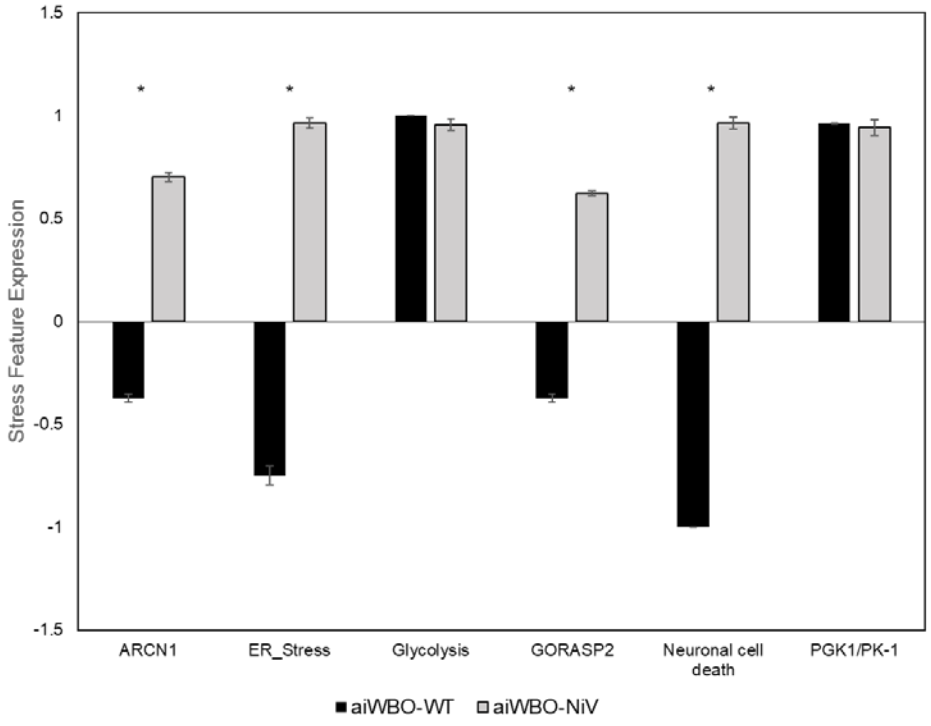


Figure 5. Cellular Stress and Neuronal Cell Death in aiWBO-NiV. comparative simulation analysis of cellular stress and neural cell death features in aiWBO–WT vs aiWBO–NiV. The vertical y-axes represent the semiquantitative levels of concepts that are estimated by NEUBOrg relative to an arbitrary baseline where 0 = baseline, 1 = maximum expression or presence and -1 = minimal expression level or presence. The horizontal x-axes represent the individual aiWBO concepts being simulated. Data represent means of three experiments ± 95% confidence interval. All p values from the Welch's t test. *$p<0.01$.

3. Summary of aiWBO–NiV and aiWBO-WT Disease Profile Simulations

The aiWBO–NiV and aiWBO–WT simulations accurately predicted the expression of 44 elements consistent with a pattern seen in a whole-brain organoid, and seven elements

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

consistent with the expression of a complex metabolic and respiratory acidosis. The ai-WBO–NiV simulations also produced some changes that are unique to NiV pathogenesis when compared with the aiWBO–WT simulation features. The probability that the expression of all (N = 51) concepts were predicted by chance alone is <0.00000001 (binomial test).

4. Evaluating Potential Vaccine Candidates for Mitigating NiV virus Encephalitis

The nine NiV genome derived proteins plus placebo were used to generate 131 unique single, double and triple immunogenic protein candidates. The final ranking included 22 double and triple protein candidates but no single protein candidates. Overall, the second ranked candidate combined proteins G and L. Just four of the final candidates contained two protein targets. All four of these contained either F or G proteins and the third ranked option overall included both proteins F and G. Eighteen of the 22 candidates contained three proteins. The number 1 candidate overall combined F, G and L proteins. Further analysis revealed that the F, G, and N proteins occurred most frequently in the remaining candidates. The protein occurrence frequencies in the final 22 combinations are summarized in Table 3. Using the recent paper based on the NipahVR (http://bioinfo.imtech.res.in/manojk/nipahvr/) database [33] the potential epitopes associated with each NiV protein are also included in the last row of Table 3 and Table 4.

Table 3. ViR protein occurrences and epitope frequencies in the top 22 candidates.

Candidate Proteins	N*	P	W	V	C	M	F*	G*	L
Single N=0	0	0	0	0	0	0	0	0	0
Double N=4	0	0	1	0	0	0	2	3	2
Triple N=18	8	3	4	4	5	5	12	10	3
Epitopes [33]	**3**	1+**	**	**	3	3	**3**	**3**	**3**

F, G , L overall top choice, *Most common components of triprotein candidates, **P/W/V = 2, W/V = 1. Importantly, based on the data from [33], the above analysis reduces the number of NiV proteins of interest from 9 to 3 (by 66.7%) and the number of potential epitopes of interest from 22 to 9 (by 59.1%).

Table 4. Table of proposed epitopes and alleles for focused vaccine research and development.

Proteins	Epitopes	Alleles
L	QPSDDKRLS	HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01
L	SDSCIIHMR	HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01
L	VSMIEPLVL	HLA-A*0203, HLA-B*3501, HLA-B*51, HLA-B*0702, HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01
F	EAMKNADNI	HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01, HLA-A*6802
F	RFALSNGVL	HLA-B*3501, HLA-B*51, HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01
F	ANCISVTCQ	HLA-A1, HLA-B*5102, HLA-B*5103, HLA-B*5401, HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01
G	KINEGLLDS	HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2022 doi:10.20944/preprints202210.0217.v1

G	ILSAFNTVI	HLA-A11, HLA-A3, HLA-A*0301, HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01
G	QRIIGVGEV	HLA-A31, HLA-B*5301, HLA-B*5401, HLA-B*51, HLA-B8, HLA-Cw*0401, H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Ld, HLA-G, H-2Qa, Mamu-A*01

Discussion

A primary purpose of the current project was to extend our previous research [AA, BB] by upgrading and literature validating a computer simulation of aiWBO using the latest version of the NEUBorg, DeepNEU (v6.6) machine learning platform. We first simulated aiWBO–WT and then a NIPAH virus (NiV) infection of the brain (i.e. encephalitis). Encephalitis is the most common sequelae of NiV infection and the most common cause of death. Our unsupervised machine learning platform for simulating aiWBOs relies on the recognized ability of iPSCs to be reprogrammed and differentiated into whole-brain organoids if optimal environmental conditions, nutrients, and sufficient time are provided. These data from the current experiments confirm that the simulations successfully reproduced a profile of Nipah encephalitis consistent with a profile derived from the peer-reviewed literature (Figure 1-3). Also, the aiWBO simulations accurately reproduced all important anatomical and regional aspects of the neonatal brain. As we reported previously, [8, 9] the current simulations have also evolved the basic components of a functioning BBB and microcirculation (Figure 4). However, the physiological impact of both elements remains limited by the absence of a working systemic cardiovascular system in our current model. Importantly, these advanced aiWBO simulations employed an efficient, cost-effective, and reproducible methodology for disease modeling.

We also compared the acid-base status of aiWBO-WT with that of aiWBO-NiV simulations (Figure 4C). In summary, there was an increased and complex combined metabolic and respiratory acidosis seen in aiWBO–NiV compared with aiWBO–WT. Additionally, there was a marked increase in the intracellular H+ ion concentration [iH+] compared to the aiWBO-WT state. Both aiWBO–WT and aiWBO–NiV appeared to be adequately oxygenated in the setting of a simulated normal ambient oxygen level (21%) and the observed hypercarbia [CO2] is most likely explained by the ambient CO2 concentration which was maintained at 5% during all simulations.

Once the unsupervised aiWBO simulations (Figure 1-3) were validated against the peer-reviewed literature, the NEUBorg platform was then used to generate the aiWBO–NiV simulations. An 11-feature literature-based genotypic and phenotypic Nipah disease profile was used to assess the accuracy of simulation outputs (see Table 2). A comparison of the aiWBO simulation predictions with those of the aiWBO-NiV revealed that all simulations correctly reproduced the feature profile seen in wet lab whole-brain organoids. Importantly, the NiV simulations predicted the presence of BBB dysfunction and vasculitis (Figure 4B) in the presence of a high NiV viral load, which are considered important markers for NiV encephalitis [35]. To our knowledge, this is the first time that whole-brain organoid simulation of NiV encephalitis has been successfully attempted.

While Nipah virus infection is known to be widespread in the brain, at least three important cell types may be more selectively affected (Figure 2A). The majority of NiV has been detected in the neurons and microglia of the brainstem, cerebral cortices, and cerebella (Figures 2 and 3A). The brain vascular endothelium is the most likely initial target of the NiV virus during the acute infection (Figure 2A) [35]. The results from our current simulations confirm that NiV virus infection is widespread within the brain with clear evidence of involvement of the vascular endothelium and other markers of encephalitis.

While there are currently no effective therapies for human Nipah disease including vaccines, the HeV-sG-V Nipah vaccine candidate has reached phase 1 human trials to assess its' safety and immunogenicity (the ClinicalTrials.gov; Identifier, NCT04199169). While NiV virus particles (NiV-VLP) composed of F, G, M and F, V, N have proven to be

 doi:10.20944/preprints202210.0217.v1

immunogenic in animal models (LLL) we could not find data regarding the specific combinations of F,G,L or F,G,N that are recommended in the current study.

The search for optimal multi-protein vaccine candidates is complicated because there are at least eight reported immunogenic proteins and a growing number of epitopes for each protein (Table 4). The goal of the current project was to simplify this search by discovering effective viral protein(s)/epitope(s) candidates and optimizing proteins/epitopes combinations.

After simulating the protein targets (N=9) derived from the NiV genome (proteins N, P, W, V, C, M, F, G, and L) we evaluated single and combination protein or epitopes as potential immunogenic components of monovalent and/or polyvalent vaccine candidates as discussed previously in [2, 28-34]. We included placebo, untreated aiWBO-NiV simulations for comparison.

The NEUBorg analyses generated 131 unique single, double, and triple immunogenic protein candidates. We have used our validated distance measure, the ACD (see the method section for more details), and evaluated all the 131 treatment options for their efficacy in preventing the Nipah encephalitis (aiWBO-NiV) and to maintain placebo (aiWBO-WT) profile in the presence of Nipah infection. Our simulations narrowed down the immunogenic protein candidates to 22 double and triple protein candidates, but no single protein candidates appeared to be effective (Table 3). The statistical significance of differences between aiWBO-NiV/placebo and treatment effects were ranked using the non-parametric Wilcoxon test and a p value < 0.05 was interpreted to indicate that observed differences were unlikely to be due to chance alone.

Overall, our data indicate that the number 1 ranked Nipah vaccine candidate is achieved by combining F, G, and L proteins. Further analysis revealed that the F, G, and N proteins occurred most frequently in the remaining candidates. The second-ranked candidate combined proteins G and L. Just four of the final candidates contained two protein targets. All four of these contained either F or G proteins and the third-ranked option overall included both proteins F and G. Eighteen of the 22 candidates contained three proteins. Importantly, based on the data from [33], the above analysis reduces the number of NiV proteins of interest from 9 to 3 (by 66.7%) and the number of potential epitopes of interest from 22 to 9 (by 59.1%). Additionally, our data also suggest that the use of viral multi-epitopes design (Table 4) as our simulation predicted that they are equally effective in designing Nipah vaccine compared to their protein counterpart (Table 3). Using viral multi-epitopes vaccine candidates have also been recommended for designing other vaccines including SARS-CoV-2 [36].

Conclusions and Future directions

Therefore, our data suggest that combining F, G, and L proteins/multi-epitope chimeras would theoretically provide optimal lead candidates for Nipah vaccine development. Additionally, this vaccine candidate discovery approach could be quickly applied, with focused wet-lab experimentation, in the event of Nipah epidemic or pandemic. Our study demonstrates that artificial intelligence-guided antigen design can be used to rapidly develop vaccine candidates for pandemic threats and supports the use of pathogen-free disease modeling as an approach to pandemic preparedness and rapid response.

CONFLICTS OF INTREST: SE and WRD have uncompensated relationships with 123Genetix medical enterprise, nonetheless the authors declare that they are providing an unbiased scientific article.

AUTHOR CONTRIBUTIONS: SE and WRD conceptualized, planned, and analyzed the experimental work, wrote the manuscript, and prepared the figures. WRD performed all computational simulation and Nipah disease modeling and vaccine identification.

REFERENCES

Villa, S., *et al.* (2020) The COVID-19 pandemic preparedness or lack thereof: from China to Italy. *Global Health & Medicine*

Singh, R.K., *et al.* (2019) Nipah virus: epidemiology, pathology, immunobiology and advances in diagnosis, vaccine designing and control strategies–a comprehensive review. *Veterinary Quarterly* 39, 26-55

 doi:10.20944/preprints202210.0217.v1

Jonkmans, N., *et al.* (2021) Scoping future outbreaks: a scoping review on the outbreak prediction of the WHO Blueprint list of priority diseases. *BMJ global health* 6, e006623

Kapata, N., *et al.* (2020) Is Africa prepared for tackling the COVID-19 (SARS-CoV-2) epidemic. Lessons from past outbreaks, ongoing pan-African public health efforts, and implications for the future. *International Journal of Infectious Diseases* 93, 233-236

Bloom, D.E. and Cadarette, D. (2019) Infectious disease threats in the twenty-first century: strengthening the global response. *Frontiers in immunology* 10, 549

Chattu, V.K., *et al.* (2018) Nipah virus epidemic in southern India and emphasizing "One Health" approach to ensure global health security. *Journal of family medicine and primary care* 7, 275

Majee, P., *et al.* (2020) Identification and characterization of two conserved G-quadruplex forming motifs in the Nipah virus genome and their interaction with G-quadruplex specific ligands. *Scientific reports* 10, 1-12

Esmail, S. and Danter, W.R. (2021) NEUBOrg: Artificially Induced Pluripotent Stem Cell-Derived Brain Organoid to Model and Study Genetics of Alzheimer's Disease Progression. *Frontiers in aging neuroscience* 13, 53

Esmail, S. and Danter, W.R. (2021) Artificially Induced Pluripotent Stem Cell-Derived Whole-Brain Organoid for Modelling the Pathophysiology of Metachromatic Leukodystrophy and Drug Repurposing. *Biomedicines* 9, 440

Takahashi, K., *et al.* (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *cell* 131, 861-872

Lancaster, M.A., *et al.* (2013) Cerebral organoids model human brain development and microcephaly. *Nature* 501, 373-379

Lancaster, M.A. and Knoblich, J.A. (2014) Generation of cerebral organoids from human pluripotent stem cells. *Nature protocols* 9, 2329-2340

Lancaster, M.A. and Huch, M. (2019) Disease modelling in human organoids. *Disease models & mechanisms* 12, dmm039347

Kim, J., *et al.* (2020) Human organoids: model systems for human biology and medicine. *Nature Reviews Molecular Cell Biology* 21, 571-584

Logan, S., *et al.* (2019) Studying human neurological disorders using induced pluripotent stem cells: from 2D monolayer to 3D organoid and blood brain barrier models. *Comprehensive Physiology* 9, 565

Chen, H.I., *et al.* (2019) Applications of human brain organoids to clinical problems. *Developmental Dynamics* 248, 53-64

Wang, H. (2018) Modeling neurological diseases with human brain organoids. *Frontiers in synaptic neuroscience* 10, 15

Chen, A., *et al.* (2020) Application of fused organoid models to study human brain development and neural disorders. *Frontiers in Cellular Neuroscience* 14, 133

Qian, X., *et al.* (2018) Generation of human brain region–specific organoids using a miniaturized spinning bioreactor. *Nature protocols* 13, 565-580

Bhaduri, A., *et al.* (2020) Cell stress in cortical organoids impairs molecular subtype specification. *Nature* 578, 142-148

Logan, S., *et al.* (2020) Dynamic characterization of structural, molecular, and electrophysiological phenotypes of human-induced pluripotent stem cell-derived cerebral organoids, and comparison with fetal and adult gene profiles. *Cells* 9, 1301

Papaspyropoulos, A., *et al.* (2020) Modeling and targeting Alzheimer's disease with organoids. *Frontiers in pharmacology* 11, 396

Kim, H., *et al.* (2019) Pluripotent stem cell-derived cerebral organoids reveal human oligodendrogenesis with dorsal and ventral origins. *Stem cell reports* 12, 890-905

Espuny-Camacho, I., *et al.* (2018) Human pluripotent stem-cell-derived cortical neurons integrate functionally into the lesioned adult murine visual cortex in an area-specific way. *Cell reports* 23, 2732-2743

Hong, Y.J. and Do, J.T. (2019) Neural lineage differentiation from pluripotent stem cells to mimic human brain tissues. *Frontiers in bioengineering and biotechnology* 7, 400

Velasco, S., *et al.* (2019) Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* 570, 523-527

McKenzie, A.T., *et al.* (2018) Brain cell type specific gene expression and co-expression network architectures. *Scientific reports* 8, 1-19

Pelissier, R., *et al.* (2019) Recent advances in the understanding of Nipah virus immunopathogenesis and anti-viral approaches. *F1000Research* 8

Kerry, R.G., *et al.* (2019) Nano-based approach to combat emerging viral (NIPAH virus) infection. *Nanomedicine: Nanotechnology, Biology and Medicine* 18, 196-220

Jensen, K.S., *et al.* (2018) Development of a novel real-time polymerase chain reaction assay for the quantitative detection of Nipah virus replicative viral RNA. *PLoS One* 13, e0199534

Aguilar, H.C. and Lee, B. (2011) Emerging paramyxoviruses: molecular mechanisms and antiviral strategies. *Expert reviews in molecular medicine* 13

Hauser, N., *et al.* (2021) Evolution of Nipah Virus Infection: Past, Present, and Future Considerations. *Tropical Medicine and Infectious Disease* 6, 24

Gupta, A.K., *et al.* (2020) NipahVR: a resource of multi-targeted putative therapeutics and epitopes for the Nipah virus. *Database* 2020

Ojha, R., *et al.* Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus Zoonotic infection, ACS Omega 4 (2019) 13069–13079.

Liu, J., *et al.* (2019) Nipah virus persists in the brains of nonhuman primate survivors. *JCI insight* 4

Kar, T., *et al.* (2020) A candidate multi-epitope vaccine against SARS-CoV-2. *Scientific reports* 10, 1-24